

⑨ 日本国特許庁(JP)

⑩ 特許出願公告

⑪ 特 許 公 報 (B 2)

昭61-43742

⑫ Int. Cl.<sup>4</sup>  
G 08 F 12/08  
12/12

識別記号 庁内整理番号  
B-8219-5B  
R-8219-5B  
Z-8219-5B

⑬ 公告 昭和61年(1986)9月29日

発明の数 1 (全33頁)

⑭ 発明の名称 周辺データ記憶システム

⑮ 特 願 昭58-10024

⑯ 公 開 昭58-147880

⑰ 出 願 昭58(1983)1月26日

⑱ 昭58(1983)9月2日

優先権主張 ⑲ 1982年2月24日 ⑳ 米国(U S)㉑ 351835

⑳ 発 明 者 マイケル・ハワード・ アメリカ合衆国アリゾナ州ツーソン・ノース・タスリ・ド  
ハーケング ライブ3866番地

㉒ 発 明 者 ジェラルド・エルスワ アメリカ合衆国アリゾナ州ツーソン・チャックワゴン・サ  
ース・テイラー ークル11373番地

㉓ 出 願 人 インターナショナル・ アメリカ合衆国10504ニューヨーク州アーモンク(番地な  
ビジネス・マシーン し)  
ズ・コーポレーション

㉔ 代 理 人 弁理士 頼 宮 孝一 外1名

審 査 官 鶴 谷 裕 二

1

2

## ⑳ 特許請求の範囲

1 それぞれ多数のデータ記憶領域を含むキャツ  
シユ及びバツキング記憶装置と、上記キャツシユ  
及びバツキング記憶装置の間でデータを転送する  
転送手段とを有する周辺データ記憶システムにお  
いて、上記キャツシユのデータ記憶領域の中で、  
上記バツキング記憶装置に記憶されたデータに対  
応してデータを記憶しているデータ記憶領域であ  
ることを表示するエントリイ、及びキャツシユに  
記憶されたデータが上記バツキング記憶装置に記  
憶された対応するデータから変更されているかど  
うかを示す変更表示を含むキャツシユ置換制御リ  
ストを保持する手段と、上記キャツシユ置換制御  
リストにおいて間もなく置換されるべき最初の部  
分を定査して、上記キャツシユのデータ記憶領域  
に記憶されたデータの中で変更されたデータを含  
む最初のデータ記憶領域を識別する第1の走査手  
段と、上記キャツシユ置換制御リスト及び上記第  
1走査手段へ接続され、上記最初のデータ記憶領  
域が識別された後に、上記キャツシユ置換制御リ  
ストの残りの部分を定査して、上記最初のデー  
タ記憶領域に記憶されたデータと所定の関係を有  
するデータを記憶したデータ記憶領域を識別する第

2の走査手段と、上記転送手段及び上記第1及び  
第2の走査手段へ接続され、上記識別されたデー  
タ記憶領域に記憶されたデータを、上記キャツシ  
ユから上記バツキング記憶装置へ転送されるべき  
データ・ブロックへグループ化する手段とを具備  
する周辺データ記憶システム。

## 発明の詳細な説明

## (本発明の分野)

本発明は多重レベルのデータ記憶階層に関し、  
更に具体的には、そのような階層において上位レ  
ベルから下位レベルへデータを転送することに関  
して、そのような階層を制御することに関する。

## 〔先行技術の説明〕

周辺データ記憶階層は、Edenその他による米  
国特許第3569938号によつて示唆されるように、  
みかけの記憶装置を与えるため、多年の間使用さ  
れてきた。Edenその他の教示によれば、要求時  
ページング又はリクエスト・システムにおいては、  
データを高速記憶装置(バッファ)へ貯蔵す  
ることによつて、周辺記憶システムは大容量を有  
する外観をとることができ、かつデータへの高速  
アクセスが可能となる。この高速アクセスは、バ  
ツキング・ストアによつて可能となるアクセスよ

(2)

特公 昭 61-43742

3

4

りも早い。更に、Edenその他の教示によれば、バッキング・ストアは磁気テープ・レコーダや磁気ディスク・レコーダのように保存性のデータ記憶装置であつてよく、前面ストアは磁気コア・メモリの如く排持久性のものであつてよい。メモリ技術の進歩にともなつて、前面ストアは、現在典型的には半導体型のメモリ素子から作られている。米国特許第3339704号は、他の形式の記憶階層を説明している。

記憶階層の重要な局面は、データの完全性を保証することである。即ち、中央処理ユニット(CPU)又はデータ処理装置のようなユーザから受取られたデータは、正確に、又はエラーが存在する旨の表示とともに供給ユニットへ戻されねばならない。従つて、記憶階層においては、保存性記憶装置のためにデータを高レベルから低レベルへ移動させるとともに、高レベルでデータを制限して他のデータを高速アクセスに備えて記憶させることが通常の典型的なやりかたである。米国特許第4020468号は、高レベル・ストアからバッキング・ストアへ変化したデータを複写する技術を教える。米国特許第4077059号は、所定の条件の下で複写データを戻す技術(コピーバック)を教える。そのようなコピーバック動作は記憶階層の効率的時間を消費する。即ち、コピーバックされるデータが多いと、現用ユニットによるデータのアクセスは能率が悪くなる。この問題は、米国特許第3588839号によつて部分的に解消することができる。この米国特許は、高レベル記憶ユニットから低レベル記憶装置へコピーバックされるデータを、変更されたデータに限る技術を教えている。即ち、バッキング・ストアに記憶されたデータと、前面ストアに記憶されたデータとの間に不一致がある場合にコピーバックが行なわれる。

記憶階層は種々の形態を有する。例えば、Edenその他による米国特許第3569838号によれば、1個の高速ストアがいくつかのユーザにサービスを与える。米国特許第3735360号は、各プロセッサがそれ自体の高速ストア又はキャッシュを有することが可能である技術を教える。更に、記憶階層の効率は、所定のデータをキャッシュ又は高速記憶機構へ貯蔵するために使用されるアルゴリズム及び制御機構によつて影響を受ける。米国特許第3898624号は、バッキング・ストアから前

面ストア(又はキャッシュ)へデータを取り出す時  
間変化が、現用CPUで実行されているプログラ  
ムに従つて、コンピュータ・オペレータによつて  
選択される技術を示す。このように、キャッシュ  
又は階層の上位レベルに存在するデータはCPU  
によつて必要とされるデータであり、必要とされ  
ない余分のデータはキャッシュに存在させないよ  
うにして、多くの有用なデータが高レベルの記憶  
機構に記憶されるようにすることが望まれる。こ  
れらの動作は、全く複雑となる。従つて、記憶階  
層をどの程度良好に管理するかを評価するため、  
記憶階層の評価プログラムが使用されてきた。米  
国特許第3964028号及び第4068304号は、これらの  
目的を達成するため、記憶階層の効率を監視する  
技術を開示している。それでもなお、各種の記憶  
階層において、データの完全性を確保しつつ、最  
適の効率を達成するため多くの技術的改善がなさ  
れねばならない。記憶階層に関する多くの改善  
は、現用CPUへ接続されたキャッシュ及びメイ  
ン・メモリの組合せについてなされてきた。メイ  
ン・メモリとキャッシュを組合せる原理は、最初  
Edenその他によつて示唆されたように、周辺シ  
ステムのデータを貯蔵しかつバッファリングする  
ことと関連して教示されてきた。Edenその他に  
よる教示は、ホスト・プロセッサの外部でデータ  
を貯蔵するか又はバッファリングすることに関す  
る。勿論、Edenその他の以前においても、CPU  
のために磁気テープ及びディスク・ユニットから  
のデータをバッファリング又は貯蔵するため、メイ  
ン・メモリが使用されてきた。即ち、メイン・  
メモリは、CPUの作業メモリとして使用される  
のみならず、周辺装置のバッファとしても使用さ  
れた。

ある種のデータ記憶システムはユーザ・データ  
を記憶し、また他の種のデータ記憶システムはベ  
ージング及びスワッピング・データを記憶する。  
これらのデータは、ベージングないしスワッピン  
グ・プログラム・データ・セットと関連して使用  
されることができる。このようにベージング・ス  
トアの例は、IBM2305固定ヘッド記憶装置であ  
る。この装置は、IBM社から出版された「IBM2835固定ヘッド記憶制御装置及びIBM2305  
固定ヘッド記憶装置の参照マニュアル」  
(Reference Manual for IBM2835 Storage

(3)

特公 昭 61-43742

5

6

Control and IBM2305 Fixed Head Storage Module) に説明されている。この周辺データ記憶システムは、記憶されたデータへ迅速にアクセスする磁気記憶ドラムを含む。磁気記憶ドラムの記憶容量は制限されているので、より大きな容量を必要とするより大型のホストは、IBM2305固定ヘッド記憶装置が最大の効率を発揮できるほど効率的に使用することができなかつた。足りない容量は、ディスク型の直接アクセス記憶装置によって補われた。特に、スワッピング・データ・セットが使用される場合、即ち、ホストと周辺メモリとの間で大量のシーケンシャル・データ・セットが迅速に転送される場合に、上記のような問題が生じた。効率を犠牲にすることなく、記憶容量の制限を除くため、IBM2305固定ヘッド記憶装置に代えて階層記憶装置を使用することができる。

ランダム・アクセス排持久性前面ストアとDASDバツキング・ストアとを有するデータ記憶階層において、前面ストア（キャッシュ）とDASDとの間のデータ転送は適当にグループ化されねばならない。その場合、単純にして効率的なデータ転送グループ化機構が、ホストの介入なしにそのような転送に必要な時間を減少させるために望まれる。特に、キャッシュの記憶スペースが他のデータを記憶するために必要とされる前に、キャッシュからのDASDへデータを移すことが望まれる。キャッシュからDASDへ、スペースが必要とされる前に、変更されたデータを移す技術は、1981年5月のIBM技術開示報告第23巻12号の5426頁及び5427頁に記載されている（IBM Technical Disclosure Bulletin, Volume 23, No.12, May 1981）。この文献の記事によれば、LRU（least recently used）-MRU（most recently used）リストにおいて変更されたデータが走査される。それは、前面ストア（キャッシュ）からバツキング・ストアへ、変更されたデータの複写を転送して、置換アルゴリズムが呼出された時、データがキャッシュからバツキング・ストアへ転送される必要がないようにするためである。この先行技術は、置換アルゴリズムを有効に働かせるため、LRU端部からLRU-MRUリストを走査して、キャッシュからDASDデータを転送する技術を教えている。しかし、それは、データのブロック転送を実行するため、関連したデータ

をバツチしたりグループ化したりする技術を教示していない。

〔本発明の要約〕

本発明の目的は、データ記憶階層のために、効率的なデータ・グループ化制御装置を提供することである。

本発明に従えば、キャッシュ（通常、非持久性を有する）置換オーバーレイ制御リストが維持される。そのようなリストは、LRUに基いていることが望ましいが、他の基準を使用してもよい。所定時間の間に、バツキング・ストア（保存性を有する）へ移動されるべきキャッシュ中の変更されたデータを発見するため、キャッシュ置換オーバーレイ制御リストが走査される。バツキング・ストアはDASDであつてよい。キャッシュ置換オーバーレイ制御リストは、前面ストアにおけるアドレス可能記憶スロットの全ての所定の百分率である利用可能閾値（可用閾値）を有する。キャッシュ置換オーバーレイ制御リスト上のスロットであつて可用閾値より下のスロットが走査されると、キャッシュからDASDへデータ転送が開始される。そのような転送を開始する前に、第2のより高い閾値（グループ化閾値又はバツチ閾値と呼ばれる）の下にあるキャッシュの全てのデータ・スロットが走査される。それは、キャッシュ内の所定の限界に至るまでの全てのデータについて、変更されかつバツキング・ストアの同じアクセス遅延境界内に存在するデータを決定するためである。DASDにおいて、遅延境界は、トラックの1つのシリンドラからトラックの他のシリンドラへ切換えられる。所定の限界は、キャッシュへの（及びキャッシュからの）データ転送特性によつて決定される。1度、関連したデータ・ブロックの全てが決定されると、それらデータ・ブロックは1つの待ち行列へグループ化される。次いで、データ転送が開始される。

本発明の第2の局面に従えば、放棄リストが維持される。放棄されるべきデータ・ブロックを発見するため、先ず放棄リストが検査される。データ書込転送を開始する前に、放棄されるべきブロックの全てが先ず放棄される。このようにして、キャッシュのアドレス可能記憶領域は、第1に放棄されることのできるデータに基いて、第2にバツキング・ストアに記憶されねばならないデータ

(4)

特公 昭 61-43742

7

に基いて、データ記憶のために割当可能とされるか、又は自由に（解放）される。

（実施例の説明）

ここで図面を参照する。これらの図面において、同一の番号は同一の構成部分を示す。階層周  
辺データ記憶システム10は、ホスト・プロセッサ（以下、ホストともいう）11のためにデータ  
信号を受取り、それを記憶し、又はそれを供給するため、ホスト11へ接続される。記憶システム  
10の典型的アプリケーションにおいて、ホスト11は中央処理ユニット（CPU）を含む。他の  
アプリケーションにおいて、ホスト11は1つの仮想計算機であつても、ハードウェアCPUの上  
で走る1組の仮想計算機であつてもよい。更に、ホスト11は多重プロセッサであつても、付加  
プロセッサなどを有する単一プロセッサであつてもよい。本発明は非常に多様な記憶システム10に  
適用することができるが、望ましい実施態様としては、データ・セットのページング及びスワッ  
ピングを処理するページング周辺データ記憶システムに適用する場合もある。概して、そのような  
ページング及びスワッピング・データ・セットは、ホスト11のためにプログラム・データ・セット  
を記憶することに関連している。そのような場合、記憶システム10は、通常、単一のホストへ  
付加される。しかし、一般的アプリケーションの周辺記憶システムは、複数のホストへ付加され  
ることができる。本発明は、周辺データ記憶システムのいずれの型式に対しても適用することが  
できる。

記憶システム10とホスト11との間の通信は、複数のI/O接続12-15を介して実行される。I/O接続12-15は、IBMシステム/370の入出力周辺チャンネルに従つて構成される。そのようなI/O接続（通常、システム及びサブチャンネルと呼ばれる）は周知であるから、これ以上説明しない。記憶システム10は、複数のDASD16（D0、D1と表示される）を含む低位（又はバッキング）記憶装置部分を有する。記憶システム10に関して、ホスト11によつて  
40 実行されるデータのアクセス及び記憶の全ては、DASD16をアドレスすることによつてなされる。このアドレッシングは、チャンネル指令ワード（CCW）19及びI/O接続12-15のアーキ

8

テクチャーを使用することによつて達成される。典型的には、チャンネル指令ワード19は、アドレス・バイト20を含む。アドレス・バイト20は、指令を受取るべき制御ユニット（CU）を指定する複数のビットを含む。第2群のビットは、アクセスされるべきDASD16を独特に指定する。記憶システム10において、DASD16の各々は複数の論理装置アドレスを有する。例えば、装置D0は4つのアドレスの1つによつてアドレスされることができる。そのような複数アドレッシングは、IBM2305固定ヘッド記憶装置において、制限された範囲で実行されてきた。各装置に対する論理アドレスはアドレス・バイト20の他のビットによつても表示される。本実施例においては、4つの論理アドレスの1つを指定してDASD16をアドレスするため、2つの論理装置アドレス・ビットがホスト11によつて使用されるものと仮定する。即ち、アドレス・ビット00は、DASD16に対する直接アクセスを指示する。この場合、ホスト11は、あたかも記憶システム10が階層システムではないかの如く、DASD16に関して動作する。アドレス・ビットが01、10、11に等しい時、後述するように、DASD16のみかけの効率がそれらの3つの論理装置アドレスによつて改善されるように、階層システムがアクセスされ、DASD16からデータが取出され、又はそこへデータが与えられる。

CCW19の第2のバイトは指令バイト21である。指令バイト21は、どのような機能が実行されるべきかを、記憶システム10へ指示するコードを含む。例示された指令SPPは、ページング・パラメータ・セット（モード・セット）指令である。第3のバイトは指令修飾バイト22である。指令修飾バイト22は複数の制御フィールドを有し、これらの制御フィールドは、記憶システム10に対して、指令バイト21で示された指令を実行するための各種のモードを表示する。例えば、1つのビットSEQは、記憶システム10に対して、来たるべき転送動作で転送されるべきデータがシーケンシャル・データであることを知らせる。バイト22のSEQビットがシーケンシャル・データであることを示す時、追加の指令修飾バイト（図示されず）がCCW19に含まれる。この指令修飾バイトは、DASD16からホスト1

(8)

特公 昭 61-43742

9

10

1へ（又は、この逆方向へ）シーケンシャルなデータ・セットとして転送されるべきデータ・ブロック又はデータ・セグメントの数を示す。そのようなシーケンシャルなデータ・セットは、ページ環境においてしばしばスワッピング・データ・セットと呼ばれる。更に、バイト22はRDビットによつて「読出し及び放棄」を指定することができる。これは、1度ホスト11が階層システムからデータを取り出した後は、そのデータはキャッシュから放棄されることができるとを意味する。DASD16にあるデータは保存される。更に、他の制御ビットを使用することができる。

データ記憶階層は、半導体ランダム・アクセス型のシステム・ストレージ30を含む。システム・ストレージ30はDASD16のためのキャッシュ40を有する。キャッシング（貯蔵）原理は周知であるから、DASD16に対するキャッシュ40の目的及び意味は詳細に説明しない。制御機構31はホスト11から周辺指令を受取り、かつ論理装置アドレスの1つを使用してDASD16へアクセスするとともに、他の3つの論理装置アドレスに基いてキャッシュ40とアクセスする。データは、キャッシュ40とDASD16との間で、ホスト動作に関して非同期的にかつ自動的に記憶システム10によつて転送される。その転送は、ホスト11とDASD16との間のデータ転送と同じ原理に基いて達成される。例えば、ホスト11は、直接モードにおいて、チャネル・アダプタ32（個別的にCAA、CAB、CAC、CADと示される）、バス70、直接アクセス・コントロール（DAC）56によつて制御されるデータ回路33、装置アダプタ34、及び装置制御アタッチメント（DCA）35を介してDASD16へアクセスする。受取られたCCW19は制御機構31によつて解釈され、ホスト11とDASD16との間のデータ・フローの方向が決定されるとともに、この種の周辺データ記憶装置を制御するための他の機能が決定される。このような機能は周知である。キャッシュ40とDASD16との間の関係は、実質的にホスト11とDASD16との間の関係に等しい。即ち、ホスト11は一直のCCWによつて制御機能を実行するが、制御機構31は複数の内部制御ワード（ICW）を使用して、キャッシュ40とDASD16との間のアクセスを実現

する。ICWは、後に説明するように、CCWと同じような構成を有する。CCW19に関してICW24を変更することによつて、データ転送動作の或る程度の効率性が達成される。制御機構31はキャッシュ・アクセス・コントロール（CAC）61を含む。CAC61はシステム・ストレージ30を動作させ、ICW24を使用しかつ直接アクセス・コントロール（DAC）56を介してDASD16にアクセスする。チャネル・アダプタ32に代つて、リンケージ・ポート（LKP）・レジスタ25がCAC61とDAC56との間の転送を実現する。LKPレジスタ25については、後に第3図に関して説明する。

ICW24は、CCWの指令バイト21に対応する指令バイト26を含む。同じ指令に対するコードは同一であることに注意すべきである。追加の指令が与えられてもよく、またCCW指令の或るものは無くてもよい。指令修飾バイト27は連鎖制御ビットCを含む。このビットは、通常、チャネル・アダプタ32を介してホスト11から制御機構31へ与えられる連鎖表示と置換される。ホスト11によつて与えられる連鎖表示は、抑制アウト・タグ信号である。ICW24のアドレス・バイト23は、DASD16のアドレスの記憶されたロケーションを指定する。ICWでは、論理アドレスは使用されない。事実として、制御機構31は、記憶階層へ与えられた論理アドレスの全てを、実際の装置アドレス・ビットへ変換する。アドレス・バイト28は、装置アドレスの記憶されたロケーションを指定するのみならず、シリンダ・アドレス（C）、ヘッド・アドレス（トラック・アドレス）（H）、レコード・アドレス（R）を指定する。ブロック・アドレスの中にあるデータ・レコードは、大部分のディスク記憶装置をアドレスする時に使用されるセクタ・アドレスに対応する。実施例において、1本のトラック（Hアドレス）の上に4つのレコードが与えられる。従つて、レコード・アドレスは1、2、3、4であり、これらは基準の回転地点に関してディスクの0°、90°、180°、270°の方向に対応する。設計パラメータによつて、直交する方向とは異つた実際の回転方向を指示するようにしてもよい。受取られたレコード・アドレスは、全て4つの回転レコード・アドレスの1つへ変換される。

(6)

特公 昭 61-43742

11

キャツシュ40は、チャネル・アダプタ32及びバス41を介してホスト11との間でデータ信号を転送する。同様に、データ信号は、DASD16とキャツシュ40との間で、データ回路33及びバス42を介して転送される。キャツシュ40とホスト11又はDASD16との間で同時的転送が望まれないならば、バス41及び42が1本のバスへ結合され、このバスがデータ転送によつて時分割される。数メガバイトを含む比較的大きなメモリであるようなキャツシュ40にアクセスする場合、キャツシュ・アクセス・コントロール61が、バス64を介してハツシュ回路44へ、装置アドレス、シリンド・アドレス及びレコード・アドレスCHRを送る必要がある。マイクロコードによつて機能を実行されてよいハツシュ回路44は、DASDアドレスをハツシュ・クラス・インデキータへ変換する。キャツシュ40の記憶容量はDASD16よりはるかに小さいから、アクセスを容易にするため、DASD16のアドレス範囲は、ハツシュ・クラスと呼ばれるクラスへ集中させられる。分散指標テーブル(SIT)45は、ハツシュ回路44によつて限定されるクラスの各々について1つのレジスタを有する。分散指標テーブル45にあるレジスタの内容は、登録簿(DIR)43に対するアドレス・ポイントである。登録簿43は、DASD16をアクセスするために使用されるアドレスDCHRを含む。データがキャツシュ40に記憶される時、DASD16へDCHRアドレスとキャツシュ40のアドレスとが、登録簿43のエントリ・レジスタに記憶される。複数の装置アドレスが1つのハツシュ・クラスに対応するから、ハツシュ法を用いてキャツシュ40を走査するには、所与のハツシュ・クラス内のエントリ・レジスタを走査すればよいように、単一的にリンクされたハツシュ・クラス・リストが登録簿43のエントリ・レジスタ中に設けられる。登録簿43の内容に基づき、キャツシュ40は既知の手法を用いてアクセスされる。もし関連したエントリが登録簿43の中に発見されなければ、キャツシュ・ミスが生じる。これは、キャツシュ・アクセス・コントロール61によつて、ホスト11からデータを受取るためキャツシュ40の中でスペースを割当てるか、ICW24及びリンケージ・ポート・レジスタ25を用

12

いてDASD16からデータを転送することを必要とする。

更に、制御機構31は、ホストへ接続される制御ユニットの通常の制御部分を含む。例えば、「アドレス及び指令」評価器(ACE)50は、ホスト11から指令信号を受取りかつ状況信号をホスト11へ与えるため、バス81、52、53、54を介してチャネル・アダプタ32と通信する。「アドレス及び指令」評価器50はCCW19を評価し、かつ命令された機能を実行するよう記憶システム10へ命令を与えると同時に、連鎖条件を表示し、また記憶システムの他の部分から状況信号を受取つてそれをホスト11へ送る。直接モードでは、「アドレス及び指令」評価器50は、バス55を介して指令信号を直接アクセス・コントロール56へ与える。従つて、データ信号は、既知のDASD周辺記憶装置手法を使用することにより、データ回路33と適当なチャネル・アダプタ32との間で、バス70を介して転送されることができる。直接アクセス・コントロール(DAC)56がその機能を実行する時、DAC56は通常的方式に従つてデータ回路33を制御する。

データ記憶装置の動作は階層的に実行される。即ち、シーケンシャル・データ・セットは、最小サイズのキャツシュ及び最小の割当制御動作を使用してキャツシュ40の中に置かれる。また、データの順次性は効率的な方法で維持され、かつホスト11の動作要件を満足させるため、十分な数のデータ・ブロックがキャツシュ40の中に維持される。「アドレス及び指令」評価器(ACE)50が、アドレス・バイト20の中にある論理装置アドレスを受取つた時、記憶階層へのアクセスを指示するため、ACE50は、論理アドレスに従つて、3つのバス60の1つを介して、受取つた指令信号をキャツシュ・アクセス・コントロール(CAC)61へ与える。3つのバスは、キャツシュ40のそれぞれのアクセスを指示する論理バスである。CAC61は、受取つた指令及び修飾データを論理装置制御ブロック(LDCB)レジスタ62に記憶する。装置の各々について3つの論理装置アドレスがあることを想起されたい。従つて、もし8個のDASD16があれば、24個のLDCBレジスタ62がある。

(7)

特公 昭 61-43742

13

各論理装置の識別情報及び動作状況は、論理装置制御ブロック・レジスタ62の1つに保存される。論理装置へのアクセス（これは、キャッシュ40の中のレジスタを装置アドレスへ割当てることによって表わされる）は、ハッシュ回路44へ至るアドレス・バス64を介して行なわれる。シーケンシャル・データ及びDASD16に対するシーケンシャル・アドレス（CHR部分）が処理される或る場合には、分散指標テーブル45にある連続したレジスタをアクセスすることができる。従つて、キャッシュ・アクセス・コントロール61は、ハッシュ回路44における遅延を避けるため、バス65を介して分散指標テーブル45にアクセスする。この動作は、シーケンシャル・データが処理されている時、ホスト11に対する記憶システム10の応答を早める。

キャッシュ・アクセス・コントロール61が、読出リクエストについて、登録簿43のハッシュ・クラスを探索した結果、ミス表示信号を受取ると、DASD16からキャッシュ40へのデータ転送リクエストが、バス66及びリンケージ・ポート・レジスタ25を介して直接アクセス・コントロール（DAC）56へ与えられる。バス66上の信号は、そのリクエストに対してDAC56の注意を向けさせ、かつICWがリンケージ・ポート・レジスタ25を介してアドレス可能であることを教える。後に明らかになるように、実施例において、リンケージ・ポート・レジスタ25はマイクロコード・リンケージ・ポート・レジスタである。直接アクセス・コントロール58は、それがCCW19に対して応答するように、ICW24へ応答する。リンケージ・ポート・レジスタ25を介してリクエストされたデータ転送（DASD16からキャッシュ40へ）が完了すると、直接アクセス・コントロール58はバス67を介してキャッシュ・アクセス・コントロール61へ状況信号を与える。その時点で、キャッシュ40はホスト11で利用可能なデータを有する。書込ミスが起ると、通常、キャッシュ40のスペースが割当てられ、次いでホストからデータが受取られる。

更に、キャッシュ・アクセス・コントロール61と直接アクセス・コントロール58との間の通信は、バス68を介して行なわれる。このような

14

通信は、メッセージ・データをリンケージ・ポート・レジスタ25に記憶することを含む。DASD16は複数の論理装置アドレスを介してアクセスされるので、1群の待ち行列レジスタ（Q）69は、キャッシュ・アクセス・コントロール61によってリクエストされた、装置に関連した動作を待ち行列として含む。このようにして、直接アクセス・コントロール58は、論理装置アドレスによるリクエストに関与せず、ホスト11又はキャッシュ・アクセス・コントロール61のために、直接アクセスDASDモードで動作することができる。このようにして、直接アクセス・コントロール58は階層システムと関連して使用されるのみならず、階層構成をとらない周辺記憶システムの中で使用されることことができる。

更に、キャッシュ・アクセス・コントロール（CAC）61は、他の制御機構を含む。例えば、ADEBレジスタ76は、CAC61が現在一緒に動作している登録簿43の1つのエントリを含む。キャッシュ40のヒットを生じたDASD16のアドレス、又はキャッシュ40の1部はホスト11によって与えられるデータへ割当てられる。レジスタ76にエントリを置くことによって、キャッシュ・アクセス・コントロール61の動作が改変される。即ち、登録簿43はシステム・ストレージ30の1部であり、レジスタ76にアクティブなエントリを入れることによって、システム・ストレージ30は解放され、制御機構31から独立して、バス41及び42を介して、データを転送することができる。装置バッファ（DEV BUF）レジスタ77は、DASD16に関連した制御情報を含み、直接アクセス・コントロール58を介してアクセスを設定する場合に、キャッシュ・アクセス・コントロール61によって使用される。このようなレジスタは、マイクロコード実施例中の書込可能制御ストア中に設けられる。レジスタ77は単に制御ストア中の割当てられた部分であり、データ構造は指定されない。バッファ・シーケンス・テーブル（BST）78については、第3図を参照して後に説明する。BST78は、バス42を介してデータ・ブロックのシーケンスとして転送される各データ・ブロックのために、登録簿43に対するポインタを含む。また、BST78は、シーケンシャル転送の間に、

(8)

特公 昭 61-43742

15

どの登録簿指標がキャッシュ40のアクセスに使用されるべきかを決定する走査制御機構を含む。このようにして、シーケンシャル転送はアドレスの設定を必要とせず、後に明らかになるように、DASD16からのブロックが中断を生じることなく迅速に転送される。チャネル制御ブロック(CCB)レジスタ63は、チャネル・アダプタ32を介して現在実行されつつあるデータ転送についての制御情報を含む。自由ブロック・リスト(FBL)・レジスタ79は、キャッシュ40のアドレス可能記憶領域(ブロック・スロットなど)のいずれが、ホスト11又はDASD16から来るデータ信号を受取りかつ記憶するために、現在割当て可能であるかを決定する。本発明の1つの目的は、FBLレジスタ79が所与の最大数のエントリを含むようにすることがである。

登録簿43はキャッシュ置換LRUリスト46を含む。リスト46は、置換制御に備えて、キャッシュ40のアドレス可能領域のリストを配列する基準として、LRU法を使用する。リスト46の詳細は第4図に示される。説明の順序として、本発明を容易に理解させるため、リスト46は第1図の下方にも示される。リスト46は第1の部分47及び第2の部分48を有する。第1の部分47は、アドレス可能領域に含まれるデータが変更されたかどうかを示す(ビット1は変更、ビット0は変更なし、Mビット)。アドレス可能領域は、ホスト11によつてチャネル・アダプタ32を介して最も最近時に使用された(MRU)順序に従つてリストされている。キャッシュ40のアドレス可能記憶領域の中で、最も最後に参照された領域が、最も最近時に使用されたアドレス可能記憶領域である。そのような記憶領域は、LRUリスト46の中でMRU-0としてリストされる。第1図に示されるように、例えばキャッシュ・アドレスによつて指定されたMRU-0は変更されている。即ち、最後の参照動作は書き込みであつた。同様に、次の「最も最近時に使用された」アドレス可能記憶領域はMRU-1である。第1図では、MRU-1も変更されている。従つて、キャッシュ40の中のデータは、対応するDASD16に記憶されたデータとは異っている。しかし、第3の「最も最近時に使用された」アドレス可能記憶領域は、対応するDASD16に記憶

16

されたデータと同じデータをキャッシュ40に記憶されている。使用時点が新しくない記憶領域は、最も使用時点が新しくない記憶領域LRU+0に至るまで、順序に配列される。従つて、次の「最も使用時点が新しくない」記憶領域はLRU+1である。以下同様である。大型のキャッシュ40では、例えば何千という多数のエントリがLRUリスト46に存在することは注意されたい。

LRUリスト46の中で、LRU+0とLRU+Nとの間にある、キャッシュ40のアドレス可能記憶領域の全ては、可用閾値49の下にある。LRUリスト46が走査され、かつアドレス可能記憶領域が変更されたデータを有することを、エントリが表示する時、それぞれの変更されたアドレス可能記憶領域のために、キャッシュ40に記憶されたデータがDASD16へ転送され始める。第2の閾値は上位閾値(又はグループ化閾値)57である。この閾値は、データをDASD16へ単一のデータ・ストリームとして転送するため、キャッシュ40のアドレス可能記憶領域からデータをグループ化するために使用される。前述したように、このグループ化は、所与のDASD16における1つのシリンドラ(多くのトラックを含む)に限定される。シリンドラ上のトラックは、全て同じ半径位置にある。即ち、これらのトラックは、1群の変換器を半径方向に移動させることなく、電気的切換えによつてアクセスされることができる。従つて、データ転送がLRU+0と可用閾値49との間の走査によつて開始される時、DASD16へ書込まれるデータ・ブロックのグループをアセンブルするため、走査はグループ化閾値57まで続けられる。例えば、LRU+1はキャッシュ40で最も使用時点の古い、変更されたデータ・ブロックであるとする。これは、複数のブロックをDASD16へ転送させる。例えば、LRU+NはLRU+1と同じシリンドラにあるとする。従つて、LRU+Nは書き込みのためにLRU+1とグループ化される。可用閾値49の上にあるMRU-Mは異なつたシリンドラにあるとする。従つて、それはグループの外に置かれる。キャッシュ40にある他のデータ・ブロック、例えばグループ化閾値57の下にあるMRU-(K+1)は変更されており、かつLRU+1と同じシリンドラに



(9)

特公 昭 61-43742

17

あるものとする。グループ化閾値57の上にあるデータは、全てDASD16へ書込まれない。何故ならば、使用時点が新しいため、近い将来、データが再びホスト11によつて参照される可能性が強いからである。そのようなデータの参照は、書き込み参照であるかも知れない。従つて、グループ化閾値57の上にあるデータをキャッシュ40から移動させることは、不要なデータ転送となる。これは効率を低下させる。

可用閾値49及びグループ化閾値57は、記憶システム10がホスト11と共に現在動作している特定の動作特性に基いて選択される。これは、閾値49及び57が或る時間の間で変化してもよいことを意味する。大切な事は、先ず可用閾値49まで走査が実行されると、データ転送が開始されることである。次に走査がグループ化閾値57まで続けられると、DASD16へグループ・データ転送を開始するため、データ・ブロックのグループ化が行なわれる。このグループ化は、DASD16が選択されねばならない回数を減少させ、従つてデータ転送の効率を高める。

第2図は、第1図に示した記憶システムの変更例である。この変更例は制御機構31の中にプログラム化されたマイクロプロセッサ31Pを含む。バス70は、チャネル・アダプタ32からデータ回路33へ接続され、第1図の場合と同じように動作する。バス41及び42は、それぞれチャネル・アダプタ32及びデータ回路33からシステム・ストレージ30へ接続される。バス41及び42は1つのバスへ結合されてよく、その場合、データ転送は1本のバスを時分割して使用する。マイクロプロセッサ31Pは、データ回路33とシステム・ストレージ30との間でデータ転送を制御するに当つて、バス71を介してデータ回路33へ制御信号を与え、バス72を介してシステム・ストレージ30へアドレス信号及びシーケンス制御信号を与える。複数のシステム・ストレージ・アドレス・レジスタ(SSAR)58は、システム・ストレージ30へアドレスを与える。例えば、8個又は16個のSSAR58が設けられてよい。マイクロプロセッサ31Pがシステム・ストレージ30へアクセスする時、それはシステム・ストレージ30のアドレスをSSAR58へ与えるだけでなく、システム・ストレージへアクセ

18

スする場合にどのSSARを使用すべきかを指示する。1つのメモリに対して多重のアドレス・レジスタを使用することは既知であり、従つてこれ以上説明しない。SSAR58はシーケンシャル・データ・ブロックの各バーストのために設けられ、マイクロプロセッサ31Pのキャッシュ40のアドレスをSSAR58へロードすることによつて、システム・ストレージ30を起動する。従つて、アドレスは、連続したシーケンシャル・ブロックの中間で、SSAR58へロードされる必要はない。シーケンシャル転送の間、マイクロプロセッサ31Pは、キャッシュ40とDASD16との間でデータ信号の転送を開始するため、単にSSARを参照するだけでよい。ここで注意すべきは、キャッシュ40がシステム・ストレージ30の中で所定のアドレス・スペースを有することである。登録簿43は、異つた範囲のアドレスを有する。SSAR58は、システム・ストレージ30のメモリ・アレイの外に設けられた別個の電子レジスタである。マイクロプロセッサ31Pは、バス51-54を介してチャネル・アダプタ32と通信する。

マイクロプロセッサ31Pの動作は、制御ストア73に記憶されたマイクロコード・プログラムに従う。制御ストア73は書込可能であることが望ましいが、1部が書込可能であつて、プログラムを含む他の部分は読出専用であつてよい。バス74は、マイクロプロセッサ31Pを制御ストア73へ接続する。制御ストア73の中には、「アドレス及び指令」評価器50の機能を実行するACEプログラム50P、直接アクセス・コントロール55の機能を実行するDACプログラム56P、キャッシュ・アクセス・コントロール61の機能を実行するCACプログラム61P、記憶システム10の動作に必要な他のプログラムであるOPプログラム75がある。しかし、OPプログラム75は、本発明を理解するために必要ではない。プログラム50P、56P、61Pによつて記憶システム10を制御するため、マイクロプロセッサ31Pによつて使用されるデータを記憶するレジスタとしては、CCBレジスタ63、LDCBレジスタ62、待ち行列レジスタ69、ADEBレジスタ76、SITレジスタ45、バッファ・レジスタ77、LKPレジスタ25、及びBSTレジス

(10)

特公 昭 61-43742

19

タ 7 8 がある。更に後述するように、FBLレジスタ 7 9、J カウンタ 3 7、放棄リスト (DL) 3 8 が使用される。キャツシュ 4 0 が非常に大きい場合、SITレジスタ 4 5 をシステム・ストレージ 3 0 に記憶することができる。効率を向上させるためには、SITレジスタ 4 5 にあるページを貯蔵するため、1組のレジスタを制御ストア 7 3 に予約しておくことができる。

第 2 図に示された記憶システム 1 0 の動作は、第 3 図から第 1 2 図までを参照すれば、最も良く理解することができる。第 3 図から第 1 2 図までは、データ構造を詳細に示すとともに、本発明の理解に必要なマイクロコードについてのマシン動作流れ図を示す。第 3 図は、記憶システム 1 0 を動作させるため、マイクロプロセッサ 3 1 P によって使用されるデータ構造を示す。LDCBレジスタ 8 2 は制御ストア 7 3 の中でデータ信号を含む一連のレジスタであり、4つの部分より構成される。最初の部分 8 0 は、基礎データ構造 (FOUND) 部分である。部分 8 0 は、記憶システム 1 0 の機能を一般的に限定するとともにサポートする。PPARMS部分 8 1 は、後に説明するセット・ページ・パラメータ指令によって設定されたページング及びスワッピング機能を限定するパラメータに関連している。CPARMS部分 8 2 は、ホスト 1 1 によって出されたセット・セクタ・シーク・サーチID指令のような指令パラメータを含む。これらの指令は既知のディスク記憶装置周辺記憶システムと組合せて使用される。RPARMS部分 8 3 は、読出活動 (即ち、DASD 1 6 からキャツシュ 4 0 へデータ信号を転送すること) をサポートするパラメータを含む。

部分 8 0 は ODE フィールド 9 0 を含む。フィールド 9 0 は、装置終了 (DE) 信号が記憶システム 1 0 からホスト 1 1 へ送られるべきか否かを示す。CNL MASK フィールド 9 1 は、チャネル・アダプタ 3 2 のいずれが現在の指令を受取ったか、即ち、論理装置がどのチャネルと提携関係を有するかを示すビット・パターンを含む。LDADDR フィールド 9 2 は、指令とともに受取られた論理アドレスを指示するコードを含む。CMD フィールド 9 3 は、第 1 図のバイタ 2 1 から取られたコードを含む。SEQ フィールド 9 4 は、第 1 図のバイト 2 2 から取られた内容を含

20

む。CCR フィールド 9 5 は、記憶システム 1 0 によってチャネル指令再試行 (CCR) 信号がホスト 1 1 へ送られたかどうかを示す。キャツシュ・ミスが MISS フィールド 9 6 で表示されると、チャネル指令再試行信号がホスト 1 1 へ送られる。従つて、フィールド 9 6 は、キャツシュ 4 0 についてミスがいつ生じたか、記憶システム 1 0 が適当な CCR 信号を与えたかどうかを示す。CCR 信号は、単にホスト 1 1 に対して、受取られた周辺指令を実行する場合に遅延が必要であるかどうかを示す。記憶システム 1 0 が指令を実行できる状態に達すると、それはホスト 1 1 へ装置終了信号を送る。次に、ホストは前の周辺指令を再び送り、その指令が記憶システム 1 0 によって実行されることになる。

PPARMS 部分 8 1 は、バイト 2 2 にあるシーケンシャル・ビットに対応する SEQ フィールド 1 0 0 を含み、かつバイト 2 2 の RD 部分から取られた RD フィールド 1 0 1 を含む。B COUNT フィールド 1 0 2 は、シーケンシャル・データとして転送されるべきブロックの数を含み。シーケンシャル・データの各ブロックがホスト 1 1 へ転送される度に、B COUNT フィールド 1 0 2 は 1 だけ減少される。従つて、それはキャツシュ 4 0 を介してホスト 1 1 へ転送されるべきブロックの数を示す。BASE CYL フィールド 1 0 3 は、シーケンシャル・データが DASD 1 6 から転送される場合のシリンドラ・アドレス C を含む。即ち、多重シリンドラ・リクエストにおいて、BASE CYL フィールド 1 0 3 は、仮想計算機 (VM) ミニディスクの値 C を含む。

CPARMS 部分 8 2 は、SEEK ADDR フィールド 1 0 4 に DASD シーク・アドレスを含み、SID フィールド 1 0 5 に最後又は現在のサーチID引数を含み、SECTOR フィールド 1 0 6 に最後又は現在のセット・セクタ値を含む。

RPARMS 部分 8 3 は、DASD 1 6 からキャツシュ 4 0 へのデータ転送が必要であることを示す REQD フィールド 1 1 0 を含む。RIP フィールド 1 1 1 は、DASD 1 6 からキャツシュ 4 0 への読出しが進行中であることを示す。RA フィールド 1 1 2 は、DASD 1 6 からの読出しが完了したことを、またある後処理機能が実行されていることを示す。DADDR フィールド 1 1 3 は、アドレスさ

(11)

特公 昭 61-43742

21

れている実際のDASD 16を示すため、バイト 20 (第1図) からの装置アドレスを含んでいる。DIR INDEXフィールド 114は、登録簿43の指標値を含む。この指標値は、どの登録簿エントリが、特定のLDCBレジスタ62で指定された論理装置に対応するエントリを含むかを表示する。SSARフィールド 115は、DASD 16とキャッシュ40との間のデータ転送において、どのSSAR 58がキャッシュ40をアクセスするとき5に使用されるかを示す。SAVEフィールド 117は、各種の動作(割込動作を含む)の間、制御データ信号を保存するため、マイクロプロセッサ31Pが使用するLDCBレジスタ62の領域を示す。

ADEBレジスタ76の構成は、登録簿43の各エントリの構成と同じである。従つて、ADEBレジスタ76の説明は、登録簿43についてもあてはまる。登録簿43及びADEBレジスタ76の各エントリにおいて、INDEXフィールド 107は登録簿エントリの論理アドレスである。このフィールドは、各エントリのために自己識別データを含む。フィールド 108は、キャッシュ中に記憶されたデータ、又は記憶のために割当てられたデータに対応するDASD 16のアドレスを含む。CCPは物理シリンド・アドレス、即ち DASD 16のためのシリンドの実際の物理アドレスである。Hはヘッド・アドレスであり、Rはレコード・アドレスであり、Dは装置アドレスであり、SECTORは実際のセクタ値である。このセクタ値は、データ・アクセスが開始されるディスクの回転位置を示す。4つのレコードを有するトラックに対するR値は、1から4まで変化することができ、セクタ値は実際のセクタ・アドレスを示す。DASDをアドレスする場合、R値は、通常35のDASDアドレッシング手法におけるように、バイト・レベルの回転位置指示情報へ変換される。R値は、ある種のホスト・オペレーティング・システムにおいて1から120またはそれ以上の数をとることができる。120を超える場合、R値は1つのトラックにあるレコード数Nのモジュロ値へ減少させられる。次いで、モジュロNのR値は、ディスクの回転アドレスへ変換される。そのようなセクタ値は、待ち時間の遅延を最小にして、レコードへのアクセスを開始させるのに適している。

22

CLLは、物理装置上で限定される論理装置のために与えられるような、論理シリンド・アドレスである。LINKフィールド 109は、1つのハッシュ・クラスの全てのエントリを一緒に結合するため、単一的にリンクされたリストのデータ信号コードを含む。所与のハッシュ・クラスの最後のエントリは、特定のコード・パターン(ゼロ)を有する。このコード・パターンは、連鎖の終り、又はクラスの終りを示す。Mフィールド 124は、キャッシュ40にあるデータが、DASD 16から受取られた後、変更されたかどうかを示す。Mフィールド 124は第1図の部分47に対応する。MRUPフィールド 125は、キャッシュ40の次の「使用時点がより新しい」アドレス可能記憶領域を指すポイントであり、MRUP 126は、キャッシュ40の次の「使用時点がより新しくない」アドレス可能記憶領域を指すポイントである。これらのポイントは、それぞれのエントリに対するフィールド 107の指標値である。MRUPフィールド及びLRUPフィールドは、後に第4図を参照して説明する二重リンクのリストと共に使用される。それぞれの登録簿エントリに対しては、他のコードを付け加えてよいが、これらのコードは、本発明の理解に必要ではない。

LKPレジスタ25は、制御ストア73にあるプログラム50P、56P、61Pによつてアクセス可能な領域である。この領域は、これらマイクロコード・プログラムの実行の相互作用を制御するため、リンケージ・ポート又はメツセージ領域を形成する。1つの実施態様では、プログラム50P及び56Pは1つのコード・セグメントとして処理され、LKPフィールド 25は、これら2つのマイクロコード部分によつて1つのユニットとしてアクセスされる。コード・ポイント(CP) フィールド 85は、制御データをポートに置いたマイクロコード部分を指定する。プログラム61PがエントリをLKPレジスタ25に置くと、プログラム56Pはその制御データを取り出し、その機能を実行する。プログラム56Pが、プログラム61Pからのリクエストに応答して、新しいデータをLKPレジスタ25に置くと、CPフィールド 85は、連続した処理のために、プログラム56Pが応答したコード実行中の地点がど

(12)

特公 昭 61-43742

23

24

こであるかを、プログラム 61P に教える。PRIORITY フィールド 86 は、LKP フィールド 25 に置かれたリクエストが高優先順位であるか、低優先順位であるか、または連続処理を示すものであるかを表すコードを含む。V フィールド (V ビット) 87 は、LKP レジスタ 25 のエントリが有効であるかどうか (即ち、それがアクションを必要とする最近のエントリであるかどうか) を示す。DADDR フィールド 88 は、DASD 16 のどれが LKP レジスタ 25 の現在の制御データ信号に関連しているかを表すため、バイト 20 から取られた装置アドレスを含む。PARMS フィールド 89 は、メッセージ (即ち、どの機能が実行されるべきであるか、状況など) と関連した各種のパラメータを含む。

BST レジスタ 78 は、DASD 16 の各々に対する 1 組のレジスタを有する。最初のレジスタは DELEP フィールド 120 を含む。フィールド 120 は、登録簿指標フィールド 122-123 を指定する指標値 1-8 を含む。これらの指標値は、削除されるべき登録簿エントリを指定する。EK フィールド 121 は、テーブル中にある有効なエントリの数を含む。更に、それはアドレスとして使用される。例えば、最初の登録簿ポインタ指標は、常にフィールド 122 に記憶され、第 8 番目の指標は、常にフィールド 123 に記憶される。EK フィールド 121 に 3 の値があると、第 3 番目の登録簿指標がアクセスされる。登録簿指標は登録簿 43 のエントリの論理アドレスであり、従って登録簿 43 への迅速なアクセスが可能となる。

システム・ストレージ 30 と DASD 16 との間の動作は非同期的であり、かつホスト 11 と DASD 16 との間の動作、及びホスト 11 とシステム・ストレージ 30 との間の動作から実質的に独立している。そのため、読出待ち行列 128 及び書込待ち行列 127 が待ち行列レジスタ 89 に設定される。書込待ち行列 127 は、装置 D0, D1, D2, D3, D4 の各々のために別個の待ち行列を有する。書込待ち行列は、書込動作を表すエントリについて、INDEX フィールド 107 に記憶された指標を含む。更に、書込待ち行列は、DASD 16 へアクセスするために必要なアドレスを記憶している。これらのアドレスはシ

ーク引数、サーチ引数、セクタを含んでいる。更に、書込待ち行列は、キャッシュ・ブロック・アドレスやシステム・ストレージ・アドレス・レジスタ 58 のようなキャッシュ 40 をアクセスするのに必要なアドレス部分、及び装置マスクを含む。各装置について、書込待ち行列 127 にある次のエントリを指定し、または次の装置書込動作 (即ち、キャッシュ 40 中のどのレコードが次に DASD 16 へ書込まれるべきか) を指定するため、リンク・フィールドを設けることができる。従って、書込待ち行列 127 は、DASD 16 の各々へ書込まれるべきデータ・ブロックにアクセスするため、キャッシュ 40 のアドレス可能記憶領域に対する 1 つ又は複数の参照項目を含むことができる。各書込待ち行列はフラグ・カウンタ 127F を含む。フラグ・カウンタ 127F がゼロでない時、それはフラグ・カウンタ 127F にある計数値に等しいエントリが書込待ち行列 127 にあることを示し、フラグ・カウンタ 127F がゼロである時、それは関連した書込待ち行列 127 が空であることを示す。書込待ち行列カウンタ (WQK) 127K は 8 ビットのシフト・レジスタであつて、書込待ち行列 127 のいずれが次にエントリを走査されるべきかを指定するため、上記シフト・レジスタの 1 ビットがセットされる。各走査は WQK 127K を 1 ビットだけシフトし、それによつて次の書込待ち行列 127 が指定される。

更に、待ち行列レジスタ 89 は読出待ち行列 128 を含む。読出待ち行列 128 の各々はポインタ 129 を有する。ポインタ 129 はラウンド・ロビン待ち行列である。読出待ち行列 128 は、どの論理装置が DASD 16 からキャッシュ 40 へのデータ転送を必要とするかを指定する LDCB レジスタ 82 のアドレスを含む。ポインタ 129 は、どの論理装置が次にサービスされるべきかを示すため、各読出待ち行列の中にある 3 個の可能なエントリの 1 つを指定する。ゼロのポインタ 128 は、空の待ち行列を表示する。例えば、装置 D0 のポインタ 129 は 2 を含むが、これは LDCB レジスタ 2 が次にサービスされるべきことを示す。LDCB レジスタ 2 の次に LDCB レジスタ 3 がサービスされる。その次に LDCB レジスタ 1 がサービスされる。D0 は、それに関連した全て

(13)

特公 昭 61-43742

25

の論理装置について、満杯の読取待ち行列を有する。装置D1の読出待ち行列は2つのエントリを有する。それらは、LDCB1及びLDCB3である。ポインタ129の数は3であるから、これはLDCBレジスタ3が次にサービスされるべきことを示している。読出待ち行列128にアクセスすることによつて、マイクロプロセッサ31Pは、データがホスト11によつてリクエストされた順序に関して、適当なシーケンスでDASD16からキャッシュ40へデータを転送するため、正しいLDCBレジスタ62へアクセスすることができる。本発明を実施するに当つて、他の優先順位決定方式を使用することができる。第3図に示されたレジスタの外に、記憶システム10を構成するため他のレジスタを使用することができる。このようなレジスタは、本発明を理解するためには必要でないので、説明を省略する。

LRUリスト46は、第1図に示されるような後入れ先出し単一リンク型スタックではなく、第4図に示されるような二重リンク型リストであることが望ましい。第4図のリストは、登録簿43の指標に従つて構成される。この指標は、欄116に示されるようにキャッシュ・アドレスCADDRに対応する。1の値を有するCADDRは、キャッシュ40の最低メモリ・アドレスに対応し、2の値を有するCADDRは、第2のアドレス可能データ記憶領域のための、キャッシュ40のメモリ・アドレスを示す。以下同様である。欄125は、キャッシュ40の中で使用時点がより新しいアドレス可能データ記憶領域であるようなエントリを、LRUリスト46の中で指定する。第4図の例では、M+2のCADDRや、欄125のゼロによつて示されるように、キャッシュのMRUアドレス可能データ記憶領域である。同様に、欄126は、次の「使用時点が新しくない」アドレス可能データ記憶領域を指定する。使用時点が最も新しくないものは、欄128にゼロがあることによつて知られるように、K+2である。注意すべきは、マイクロプロセッサ31Pの中にある作業レジスタ（図示せず）がMRU及びLRUのアドレス可能記憶領域のアドレスを含んでおり、LRUリスト46は、LRU端部又はMRU端部（第1図参照）のいずれからでも走査できることである。MRU及びLRUのポインタ値は、通常

26

の二重リンク型リスト・パターンで設定される。例えば、CADDRがK+2のMRUPは1の値に等しい。1の値を有するCADDRを探すと、LRUPはK+2の値を有し、MRUPはMの値を有する。これは、次の「使用時点がより新しい」エレメントである。CADDRがMであるエレメントを探すと、LRUP値は1を有し、MRUP値は2つのハイフンを有する。2つのハイフンは、リストの中に追加のエントリがあることを示す（説明を簡略にするため、これ以上説明しない）。同様に、CADDRが2である場合、LRUP値は2つのハイフンであり、これはエントリが省略されていることを示す。可用閾値49及びグループ化閾値57は、LRUリスト46の中でどこにあつてもよい。そのような閾値を走査する場合、走査されたエレメントの数がカウントされ、閾値に達したかどうかを決定するため、ある数値と比較される。これについては、第10図を参照して詳細に説明する。

第5図は、読出データ転送及び書込データ転送におけるCCW及びICWのシーケンスを示す。読出転送は、DASD16からホスト11へ信号を転送し、書込転送はその反対におけるデータ転送である。CCW連鎖130は、セット・ページパラメータ（SPP）指令132で始まる。第12図は、そのような指令が記憶システム10によつて実行される有様を示す。基本的には、SPP指令132は、キャッシュ40からホスト11へ読出されたデータが放棄されることのできるかどうかを設定するとともに、CCW19（第1図）のバイト22で指定された他のパラメータを設定する。1度SPP指令が動作パラメータを記憶システム10へ指示すると、シーク指令133によつて、シーク指令が周辺記憶システムへ転送される。1つの実施例において、シーク・パラメータはSPP指令に含まれている。通常のDASDアーキテクチャに従えば、シーク指令の次にセット・セクタ指令134が来る。セット・セクタ指令の次にはサーチID等価（SIDE）指令135が来る。ここで記憶システムは、読出指令136によつて、アドレスされたDASD16からデータを読取る準備を完了する。読出指令を受取ると、記憶システム10は、スタック131で示される動作を準備する。先ず、シーク指令、セット・セクタ指令、

(14)

特公 昭 61-43742

27

ID指令が140でスタックされる。137では、第1図に関して説明したように、登録簿43のサーチが実行される。キャッシュ・ヒットの場合（即ち、リクエストされたデータがキャッシュ40にある場合、矢印138で示されるように、データは、直ちにキャッシュ40からホスト11へ、指令を受取ったチャネル・アダプタ32を介して転送される。他方、登録簿43によつて、データがキャッシュにないことが示されると、矢印141で示されるように、キャッシュ・ミスが生じる。そして、矢印142で示されるように、チャネル指令再試行（CCR）信号が記憶システム10によつて与えられる。CCR信号は、いつ装置終了信号が記憶システム10から受取られたか、及び同じ読出指令を記憶システム10へ送ることによつて、読出指令136がチャネルによつて再実行されねばならないことをホスト11へ知らせる。これが起つている間に、記憶システム10は、シークICW143で始まるICW連鎖143-148を組立てる。シークICW143は、ホスト11から受取られたスタックのシーク指令から引出される。多重トラック動作においては、ICWはサーチIDパラメータから引出される。シークICW143の次にセット・セクタICW144が来る。セット・セクタICW144は、レコード番号から計算されたセクタを有する。148において、局所入力によつてセット・キャッシュICW145が生じる。このICWによつて、プログラム50Pは、読出されるべきデータが記憶されるシステム・ストレージ30のアドレスを、適当なシステム・ストレージ・アドレス・レジスタへ入れる。もし複数のデータ・ブロックが転送されるべきであれば、146で示されるように、複数のセット・キャッシュICWが起る。次に、SIDE指令135に対応するサーチID等価ICW147が起る。サーチID等価ICW147は、最初のセット・キャッシュICW145に対応する。これは、複数のデータ・ブロックが、1つのSIDE指令147を使用して順次に読出されることを意味する。次に、転送されるべきデータ・ブロックの数に等しい読出ICW148がプログラム50Pへ与えられる。それは、セット・キャッシュICWの数によつて示されたデータ・ブロックの数を読出すためである。アドレスされた

28

DASD16から、キャッシュ40のシステム・ストレージ・アドレス・レジスタ58によつて指定されたアドレスへ、データを転送する読出動作が完了すると、記憶システム10は、ホスト11へ矢印150によつて示されるように装置終了（DE）信号を与える。ホスト11は、指令136に対応する周辺指令を151で再び出すことによつて、直ちに応答する。勿論、記憶システム10は152で登録簿43をサーチし、その結果、ICW連鎖が実行されたばかりであるから、キャッシュ・ヒットを生じる。次に、データは、矢印153で示されるように、キャッシュ40からホスト11へ転送される。136において、リクエストされたデータ・ブロックについて、データが転送されなかった場合、他のキャッシュ・ミスが生じ、エラー状況がホスト11へ報告される。このエラー状況は、記憶システム10が、DASD16のリクエストされたシリンダ及びヘッド・アドレスからデータを転送できなかった事実を反映する。その場合、ホスト11は、標準のディスク記憶装置回復手法を使用して、直接アクセスの経路からデータを転送することができる。省略符号154は、前述した動作が反復されること、及び複数のDASD16について複数のCCW連鎖が間挿されてよいことを示す。ICW連鎖は、必ずしもCCW連鎖の順序に従わなくてもよい。状況に応じて、ICW連鎖は、後に発生するCCW連鎖によつて組立てられかつ使用されてよい。そのような可能性は、CCW連鎖に関してICW連鎖の非同期的局面を示す。通常、最初のCCW連鎖によつて、最初に発生するICW連鎖を生じる。任意の時点で、別個のICW連鎖が各装置に対してアクティブであつてよい。シーク、セット・セクタ、SIDEの指令133-135は周知であるから、これ以上説明しない。

第6図はシーケンシャルな読出データを転送するためのマシン動作を示す。160において、受取られた指令は、プログラム50Pによつて処理される。次に、LKPレジスタ25を介して、プログラム61Pがマイクロプロセッサ31Pによつて能動化される。161で、指令は再びプログラム61Pによつてデコードされる。それは読出指令であるから、第1図に関して説明したようにして、登録簿43がサーチされる。163で、登

(15)

特公 昭 61-43742

29

録簿のサーチの結果、キャッシュ・ヒットが生じたか、キャッシュ・ミスが生じたかを、マイクロプロセッサ 31P が決定する。キャッシュ・ミスの場合、164 で、指令及びその制御情報を待ち行列レジスタ 89 に置くことによつて、受取られた指令が待ち行列に入れられる。また、CCR 信号がホスト 11 へ送られる。待ち行列 89 は任意の様式を有してよい。待ち行列 89 は、DASD 16 の各々に対する先入れ先出し (PIFO) 待ち行列である。即ち、8 個の DASD 16 があれば、8 10 つの待ち行列が存在する。FIFO 待ち行列を使用することによつて、所与の装置に関して、ホストに対する応答の順序が、確実にホストによつて送られた指令の順序に対応することとなる。待ち行列レジスタ 89 から、プログラム 64P は、スタック 181 を形成することによつて、アドレスされた DASD 16 からの読出しを起動する。

163 で、登録簿サーチによつてキャッシュ・ヒットが生じると、キャッシュ 40 は、170 で、適当なチャネル・アダプタ 32 を介して、自動的にデータをホスト 11 へ転送する。キャッシュからホストへの、そのような自動的転送は周知であるから、詳細に説明しない。自動的データ転送の間に、エラーが起る場合がある。従つて、エラーが検出されると、マイクロプロセッサ 31P 25 は、エラー報告分析ルーチンへ行く。概して、データ転送にエラーは生じない。171 で、データ転送が成功裡に終了すると、マイクロプロセッサ 31P は LDCB レジスタ 62 にアクセスして、RD フィールド 101 を検査する。もし RD フィールド 101 に「読出し及び放棄」が表示されてい 30 れば、173 で、マイクロプロセッサ 31P は、キャッシュ中に読出したばかりのデータ・ブロックについて、制御ストア 73 の放棄リスト 38 の中に「放棄」を表示する。放棄は、指令が実行されて 35 いない時、マイクロプロセッサ 31P がシステム・ストレージ 30 の登録簿 43 にアクセスして、そのエントリを消去することにより実行される。置換アルゴリズムが呼出される前に、データを放棄することは、キャッシュ 40 を効率的に 40 管理するのに必要な制御手順を減少させる。即ち、自由な（解放された）アドレス可能データ記憶領域は、それらが必要とされる前に確保されることになる。171 で、RD=0 であれば（読出

30

し及び放棄でない)、175 で、マイクロプロセッサ 31P は、登録簿 43 の或るフィールドによつて（図示せず）、データがキャッシュ 40 へ固定されるべきかどうかを決定する。キャッシュ 40 へのデータの固定は、登録簿 43 の固定フラグ（図示せず）がホスト 11 の指令によつて消去されるまで、そのデータをキャッシュ 40 から消去できないことを意味する。もしデータがキャッシュへ固定されるのであれば、176 で、読出されたばかりのブロックは、登録簿 43 の LRU リスト中で、「使用時点の最も新しい」(MRU) ブロックとされる。177 で、本発明と無関連の論理ステップがマイクロプロセッサ 31P によつて実行される。次に 180 で、LDCB レジスタ 62 が再びアクセスされて、SEQ フィールド 100 が検査される。もしアドレスされた DASD 16 に関してシーケンシャル・データが表示されてい 15 れば、182 で、マイクロプロセッサ 31P は LDCB レジスタ 62 の B COUNT フィールド 102 を検査する。それは、ブロック・カウント (BK) が 0 に等しいかどうか、即ち、転送されたばかりのブロックが、データのシーケンス中最後のブロックであるかどうかを調べるためである。もしそれが最後のブロックでなければ、183 25 で、ブロック・カウントが 1 だけ減少される。ステップ 180、182、183 及び論理通路 181 は、マイクロプロセッサ 31P をプログラム 50P へ導く。それは、通常の態様で、最終状況をホスト 11 へ報告するためである。

第 7 図及び第 8 図は、待ち行列レジスタ 89 の中の読出待ち行列を定査し、記憶システム 10 の内部指令である ICW 連鎖を発生する動作フローを示す。リクエストされた読出指令が待ち行列に入れられた後、マイクロプロセッサ 31P は記憶 35 システム 10 に各種の機能を実行させる。例えば、そのような機能は、チャネル・アダプタ 32 を介して受取られた指令に回答すること、読出指令を転送したチャネル・アダプタ 32 から受取られた他の指令を実行したりすることを含む。ホスト 40 からの周辺指令の受取りが休止すると、シーク指令及びセクタ・セクタ指令がマイクロプロセッサ 31P によつて DASD 16 へ送られる。キャッシュ 40 がデータをホスト 11 へ送ったり、ホスト 11 からデータを受取ったり、DASD 16 へデ

(16)

特公 昭 61-43742

31

ータを送つたり又はそこからデータを受取つている間に生じる制御活動が休止すると、マイクロプロセッサ 31P は、そのタスク指名マイクロコード（これは第2図に示されるOP75の1部である）を介して、待ち行列レジスタ69を含む作業テーブル（図示せず）を走査する。もし待ち行列レジスタ69の中の待ち行列が空であれば（即ち、読出しも書き込みも起るべきでない）、マイクロプロセッサ 31P は論理通路192をたどつて190へ戻る。もし191で、ポイント129（第3図）を走査することによつて、読出指令が待ち行列に入れられていることが検出されると、193で、読出待ち行列128からマイクロプロセッサ 31P の内部レジスタ（図示せず）へ行列のエントリが転送される。もしこの転送の間にエラーが生じると、エラー報告及び回復手順が194で開始される。待ち行列レジスタ69からのエントリが読出しが成功すると、195でLDCBレジスタ62がアクセスされ、ODEフィールド90が1へセットされる。それは、読出しが成功裡に完了した時、装置終了信号を出すべきであることを示す（第5図の矢印150を参照）。196で、本発明と無関連の機能が実行される。次に200で、アドレスされたDASDに対応するバッファ・レジスタ77（第2図）の中で、第1図のバイト27にあるビットがセットされる。それは、論理連鎖が生じること、即ちアドレスされたDASD16に対する来たるべきアクセスで、1つ以上のICWが使用されることを示す。201で、LDCBレジスタ62が再びアクセスされて、SEQフィールド100の値が検査される。シーケンシャル・データが表示されていれば、マイクロプロセッサ 31P はステップ202へ進す。それは、来たるべきICW連鎖のためのブロック・カウント（BK）を、受取られたページング・パラメータ（PA）に等しくセットするためである。

所与のICW連鎖を介して転送されることのできるブロックの最大数は、システム・ストレージ・アドレス・レジスタ（SSAR）58の数に等しい。例えば、SSARが8個の場合、転送されるブロックの数は、最大8個である。更に、遅延境界を考慮に入れなければならない。例えば、転送されるべき8個のブロックが2個のシリンダにア

32

クセスすることを必要とすれば、最初にアクセスされたシリンダにあるブロックのみが転送される。もし8個のブロックの中で、最初のシリンダに4個のブロックがあり、第2のシリンダに4個のブロックがあれば、ブロック数は4にセットされる。この方法は、一連のブロックを転送するために必要な時間を最小にし、全ての転送が電子的速度で完了することを可能にする。所与のシリンダの最初のブロックでミスが生じた場合、8個のブロックまで自動的に転送される。更に、ブロックの最大数は、B COUNTフィールド102にある数値より決して大きくない。ICWの構成は、シリンダ境界が所与のICW連鎖によつて決して交差されないようになっている。このような計算は、通常のプログラミング手法によつて実行することができ、これ以上詳細に説明しない。ステップ201で、もしシーケンシャル・データが表示されなければ、ステップ203で、転送されるべきブロックの数は、1へセットされる。これらの数値は、連鎖フラグ、DASDアドレス、及び他の装置制御データと共にバッファ・レジスタ77へ与えられる。204で、SSAR58の識別記号が0へセットされる。これは、マイクロプロセッサ 31P が、識別記号0を有するSSARにアクセスすることを意味する。

205で、第1図のCCW19を介して受取られた論理アドレスLDADDRが、DASD16の物理アドレス（DADDR）へ変換される。この動作は、論理アドレス表示ビットを削除するため、単に論理アドレスをマスクすることによつて達成される。206で、本発明と無関連の機能が実行される。論理コネクタ207は、第8図で説明されるマシン動作流れ図の論理コネクタ243から接続される。即ち、190から206までの論理ステップの全ては、準備ステップであり、次に説明するステップは、連続したデータ・ブロック転送を実行するため、ループとして反復可能なステップである。

ループ中の最初のステップ210は、キャッシュ40の中でスロット又はアドレス可能データ記憶スペースを割振る。通常の割振手順がとられる。即ち、自由リスト中のアドレス可能ユニット（スロット）が、DASD16から来た最初の信号ブロックを受取るため、アドレス可能のユニット



(17)

特公 昭 61-43742

33

又はスロットとして指定される。次に、そのスロットは自由リスト (FBLレジスタ79) から除かれ、マイクロプロセッサ31Pの中にある内部レジスタに表示される。それは、キャッシュ40の中のスロットを指定するため、登録簿43の中5のどのエントリが使用されるべきかを示す。注意すべきは、キャッシュ40にある各アドレス可能スロットのために、登録簿43の中に1つのエントリ・レジスタが設けられていることである。従つて、キャッシュ40にあるデータの実際のアドレスは、登録簿43のエントリを含むレジスタから直接に引出すことができる。

ステップ202又は203で設定されたブロック数に等しい数のスロットが割振りを試みられた後、211で、マイクロプロセッサ31Pは割振プロセスでエラーが生じたかどうかを決定する。もしエラーが生じたのであれば、DASD16からキャッシュ40へ、ブロックの全数が成功裡に転送されなかったのかも知れない。従つて、エラー条件が生じると、212で、マイクロプロセッサ31Pは、LDCBレジスタ62のSEQフィールド100を調べて、データ転送がシーケンシャル転送であるかどうかを決定する。もしそれがシーケンシャル転送でなければ、マイクロプロセッサ31Pは論理通路213をたどつて、プログラム50Pへ戻り、1つのブロックに対しスペースを利用可能とするため、置換アルゴリズムの制御を待機する。シーケンシャルなデータ転送である場合、214でマイクロプロセッサ31Pは、転送されるべき最初のブロック上でエラーが生じたかどうかを決定する。もしそれが最初のブロックであれば、マイクロプロセッサ31Pは論理通路218を介してプログラム50Pへ戻る。もし割振りのエラーが最初のブロックについて生じたのでなければ、先行するブロックのデータ転送が起25る。マイクロプロセッサ31Pは通路217をたどつてステップ220へ至る。それは、割振されない領域において転送されるべきブロックの数をICWから除去するためである。即ち、割振エラーに関連させることのできるブロックから始まる全ての所望のブロックが、ICW連鎖から除去される。

次にステップ211に戻つて、もし割振エラーが起らなかったのであれば、218で本発明と無

34

関連の機能が実行される。これらの機能は、割振りに関連しないマイクロコード論理エラーを分析することを含む。もしそのようなマイクロコード・エラーのために、スロットが割振られなかったのであれば、除去ステップ220が実行されて、DASD16からキャッシュ40へ転送されるブロックの数が減少される。エラーが生じない場合、又は除去ステップの後、マイクロプロセッサ31Pは221で本発明と無関連の論理ステップを実行する。222で、LDCBレジスタ62のSEQフィールド100が検査される。もしSEQフィールド100がゼロに等しければ (即ち、シーケンシャルでないデータ)、223で、データを受取るべきキャッシュ40中のスロットに対応する、登録簿43のエントリの指標が、RPARM部分83のフィールド114へ入れられる。シーケンシャル・データの場合、又は指標がLDCBレジスタ62へ入れられた後、224で、後にシステム・ストレージ・アドレス・レジスタ58へ挿入されるべきキャッシュ・アドレスが、LDCBレジスタ62へ挿入されたばかりの登録簿指標から発生される。この発生は、単にオフセットを登録簿指標の各々へ付け加えることによつてなされる。次に225で、SEQフィールド100がシーケンシャル・モードを示すと、ブロック・カウント (BK) が1より大きいかどうかを調べるため、マイクロプロセッサ31PはB COUNTフィールド102を検査する。次に232で、マイクロプロセッサ31Pは、転送されつつあるブロックのシーケンス中にある最初のデータ・ブロックが、現在処理されているかどうかを調べる。もし処理されていないならば、233で、第2のブロックに対する新しいキャッシュ・アドレスが与えられる。次に234で、バッファ・レジスタ77の中で、第2又は他のブロックに対応するシステム・ストレージ・アドレス・レジスタ (SSAR) 58がキャッシュ・アドレスへセットされ、フラグがセットされ、登録簿43に対するポインタがセットされ、キャッシュ・アドレスを受取るべきSSAR58が指定される。更に、バッファ・レジスタ77の中に、実行すべき他の機能を限定してよい。

ステップ225、231、232へ戻つて、論理通路226は本発明と無関連のステップ227

(18)

特公 昭 61-43742

35

へ進み、次いでマイクロプロセッサ31PはLDCBレジスタ62へ進んで、RPARMS部分83のフィールド116に発生されたキャッシュ40のアドレスを記憶する。次いでマイクロプロセッサ31Pは本発明と無関連のステップ229を実行し、論理コネクタ235を介して第8図の論理ステップへ進む。

第7図の論理コネクタ235は、第8図の論理コネクタ240へ接続される。241で、マイクロプロセッサ31Pは、第3図のEKフィールド121を増加することによつて、SSAR58へのポイントを更新する。242において、マイクロプロセッサ31Pは、キャッシュ40へ転送されるべきデータ・ブロックの全てが、キャッシュ40の中でスペースの割振りを受けたかどうかを決定する。もし受けていなければ、マイクロプロセッサ31Pは、論理コネクタ243を介して、第7図の論理コネクタ207へ戻り、他のデータ・ブロックのためにキャッシュ40のスロットを割振る。このループは、EKフィールド121が転送されるべきブロック数に等しいカウントを含むようになるまで反復される。

前記のループを完了した後、本発明と無関連の論理ステップが244で実行される。245で、読出指令がICWにセットされて、DASD16のためのデータ読出指令が表示される。250で、LDCBレジスタ62がアクセスされて、RPARMS部分81にあるSEQフィールド100がセットされているかリセットされているかが決定される。セットされていれば、251で、マイクロプロセッサ31Pは、受取られたブロック・カウントが1より大きいかどうかを決定する。もしそれが1より大きければ、ICW24の指令修飾バイト27の中にセットされる。1より大きくなければ、ステップ250又は251から、バイト27の連鎖表示をリセットすることによつて、連鎖終了表示EOCが表示される。254で、制御ストア73にあるバッファ・レジスタ77は、ICWを受取る。即ち、コード、フラグ、及び他の記憶動作表示(STOROP)が受取られる。255で、マイクロプロセッサ31Pは再びLDCBレジスタ62のSEQフィールド100を調べて、それが0に等しいかどうか(即ち、シーケンシャルでないかどうか)を決定する。もし1つの

36

ブロックだけが転送されるのであれば、マイクロプロセッサ31Pは論理通路256をたどり、構成されたばかりのICWをLKPレジスタ25を介してDACプログラム56Pへ転送するため、ステップ257を実行する。

シーケンシャル・データ転送の場合、マイクロプロセッサ31Pはステップ259からステップ260へ進み、EKフィールド121を次のエントリへ調整する。次に261で、もし残りのブロック・カウントが1より大きくなければ、257で、ICW24がLKPレジスタ25を介してDACプログラム56Pへ転送される。残っているブロック数が1より大であれば、ループ270が実行されて、連鎖中の残りのICW24が設定される。271で、「カウント、キー、データ」読出指令及び多重トラック指令がセットされる。272で、マイクロプロセッサ31Pは、シーケンシャルなブロック群の中の最後のブロックが処理されるべきか否かを決定する。もし処理されるべきでなければ、形成されつつあるICWのバイト27にある連鎖フラグが1へセットされる。処理されるべきであれば、274で、連鎖フラグをリセットすることによつて、連鎖状態の終り(EOC)が表示される。275で、形成されたばかりのICW24がバッファ・レジスタ77へ転送される。276で、キャッシュ・アドレスCADDRがバッファ・レジスタ77に記憶され、それが直ちにバースト転送のためにSSAR58へ転送されることができるようになる。277で、マイクロプロセッサ31Pは、データ・ブロックが最後のブロックであるかどうかを決定する。もし最後のブロックでなければ、ループは278で調整され、通常の制御手法を用いて、内部レジスタ内のカウントが調整される。最後のブロックであれば、ステップ257が実行される。ループが278で調整された時、ステップ271から277までが実行される。

プログラム56PがICW連鎖を受取ると、それはチャネル・アダプタ32を通して受取った連鎖指令を実行するようにして、ICW連鎖を実行する。このような動作は周知であるから、ICW連鎖の実行はこれ以上説明しない。注意すべきは、DASD16からキャッシュ40へ信号を転送する時、プログラム56PはDASD16へアドレ

(10)

特公 昭 61-43742

37

シングするのみならず、バッファ・レジスタ 77 のキャッシュ・アドレスを SSAR 58 へ与え、数ブロックのデータが単一のデータ・ストリームで転送されることができるようにする。即ち、多重トラック転送ができるようにする。転送が終ると、プログラム 56 P は、エラー表示を含む結果の状況を LKP レジスタ 25 へロードする。次にマイクロプロセッサ 31 P の動作は、プログラム 56 P から CAC プログラム 61 P へ切換えられる。

第 9 図は、キャッシュ・ミスに続いて、マイクロプロセッサ 31 P を介して実行されるマシン動作を示す。これらのマシン動作は、例えば、第 8 図のステップ 164 で起る。280 で、マイクロプロセッサ 31 P はアドレスされた論理装置に対する LDCB レジスタ 62 にアクセスし、FOUND 部分 80 の CCR フィールド 95 及び MISS フィールド 96 を 1 へセットする。その動作は、キャッシュ・ミスが生じたため、チャネル指令再試行信号がホスト 11 へ送られたことを示す。これについて、後に第 11 図を参照して説明する。次に 281 で、マイクロプロセッサ 31 P は、FBL 4 (自由ブロック・リスト) レジスタ 79 を検査して、キャッシュ 40 のアドレス可能データ記憶領域又はブロックが割振りのために解放されているかどうかを決定する。キャッシュ 40 から DASD 16 へ、データを早期に転送しておれば、可能性としていくつかのブロックが解放されているであろう。FBL レジスタは、LRU リスト 46 の部分 47 がゼロに等しいことを必要とするように構成するのが望ましい。LRU リスト 46 の中に解放されたブロックを表示することによって、指定されたデータのアドレス可能性が維持される。ステップ 301-307 はこの動作を示す。解放されたブロックを発見すると、マイクロプロセッサ 31 P はステップ 284 への論理通路 282 をとる。ステップ 284 では、FBL レジスタ 79 から解放された自由ブロックの表示が除去され、登録簿 43 の対応するエントリが現在の識別表示から変更され、その時キャッシュ 40 へ割振られつつある DASD 16 のアドレス可能記憶領域が表示される。そのような割振動作の間にエラーが発生すると、マイクロプロセッサ 31 P はステップ 284 から論理通路 285 をたどつてエラー回復

38

手順へ進む。別個の FBL レジスタ 79 が設けられている場合、割振ステップ 284 は、第 11 図を参照して説明するように、論理コネクタ 283 を介してデータ転送動作に続いて実行されてよい。

284 で、キャッシュ 40 のアドレス可能データ記憶領域の割振りが成功すると、286 で、マイクロプロセッサ 31 P は LDCB レジスタ 62 の CMD フィールド 93 にアクセスし、キャッシュ・ミスがチャネル読出型の指令から生じたのか、チャネル書込型指令から生じたのかを決定する。チャネル書込型指令の場合、287 で、装置終了 (DE) 信号がホスト 11 へ与えられる。即ち、書込指令に対するキャッシュ・ミスを処理する場合、必要なことは、ホスト 11 からデータを受取るため、キャッシュ 40 の中でスペースを割振ることである。マイクロプロセッサ 31 P が十分に早い時、書込指令のミスに対して CCR 信号を送らなくてもよい。その場合、ステップ 284 及び 286 がステップ 280 に先行する。次にステップ 280 が、ステップ 291 の前にあるように、キャッシュ読出ミスのサーチに対してのみ実行される。286 において、読出動作が表示されると、291 で、マイクロプロセッサ 31 P は関連した DASD 16 のために読出待ち行列 (第 3 図) をセットする。例えば、装置 D2 の場合、それに関連した LDCB レジスタ 2 に対するキャッシュ・ミスは、読出待ち行列 128 に図示されたエントリを生じる。マイクロプロセッサ 31 P は、読出待ち行列を設定した後に、292 の動作へ進む。292 の動作は、記憶システム 10 の中で実行されるべき作業をサーチするタスク指令ステップ 190 を実行することを含む。ステップ 190 を実行した結果として、実行すべくスケジュールされた読出動作を決定するため、読出待ち行列 128 が定査される。次に、293 で、マイクロプロセッサ 31 P は、リクエストされたデータ・ブロックを DASD 16 (例えば装置 D2) からキャッシュ 40 へ転送する。勿論、これは、マイクロプロセッサ 31 P が読出待ち行列 128 の情報を、プログラム 56 P によって使用される LKP レジスタ 25 へ与えることを含む。1度、プログラム 56 P が ICW 連鎖から必要な情報を受取ると、データは、既知の DASD 読出手法を用

(20)

特公 昭 61-43742

39

40

いて、アドレスされたDASD 16からキャッシュ  
40へ転送される。その読出動作が終了すると、  
294で、マイクロプロセッサ31Pは、装置D  
2に対するLDCBレジスタ62のCCRフィールド  
95及びMISSフィールド96を検査する。その  
検査によつて、マイクロプロセッサ31Pは、  
DASD 16からキャッシュ40への読出動作がキャ  
ッシュ・ミスから生じたことを確認する。これ  
は、294で示されるように、装置終了信号がホ  
スト11へ与えられることを必要とする。

キャッシュ・ミスを生じたチャネル指令の再実行  
は、少し遅れて起る。即ち、ホスト11は、チャ  
ネル指令を再び出すことによつて、ステップ2  
87又は294で送られた装置終了信号に回答す  
る。程なく、288で、記憶システム10は、ホ  
スト11によつてリクエストされた全ての種類の  
データ記憶動作及び非同期的動作を実行するた  
め、本発明と無関連の指令を受取る。290で、  
ホスト11は第2回目の指令を与え、登録簿43  
をサーチしている記憶システム10は、今や第6  
図のステップ163で生じるようなキャッシュ・  
ヒットを有し、その結果、データはキャッシュ4  
0からホスト11へ転送される。2回目に出され  
た指令が完了すると、記憶システム10は、タスク  
指名ステップ190を介して、他のデータ動作  
へ進む。

FBLレジスタ79によつて、キャッシュ40  
に解放されたブロックがないか、又はLRUリス  
ト46を走査した結果、部分47がゼロであるエ  
ントリがない場合、キャッシュ40のアドレス  
可能記憶領域のあるものが解放されて、割振ス  
テップ284の実行を可能にしなければならない。  
そのような解放はキャッシュ・ミスの結果だけ  
でなく、タスク指令ステップ190の動作が作業を  
走査している時にも生じる。例えば、所定の優先  
順位に従つて、ステップ190はマイクロプロセ  
ッサ31Pをして論理通路299をたどらしめ、  
定期的又は非定期的に、次に説明するループ30  
0のマシン動作を実行させる。それは、FBLレ  
ジスタ79が、キャッシュ・ミスから生じた割振  
要件を満足させるエントリを有するようにさせ  
るためである。その場合、FBLレジスタ79が  
空でないと、キャッシュ・ミスをする書込指令  
について、割振ステップ284を迅速に実行さ

せ、前述したようにホスト11へチャネル指令  
再実行(CCR)信号を送るのをやめることがで  
きる。CCR信号の送信中止は、記憶システム1  
0の望ましい動作モードである。

5 本発明の場合、キャッシュ40からDASD 16  
へのデータ転送動作(以下、カースト・アウト動  
作ともいう)では、転送されるデータが指定され  
るのみならず、キャッシュ40のいくつかのアド  
レス記憶領域からの複数のデータ・ブロックが、  
10 グループとしてまとめられる。DASD 16が、電  
氣的切換えにより複数の変換器によつてアクセス  
可能な複数のシリンドを有する場合、シリンドに  
記憶可能なデータ・ブロックの全て、及び最初に  
指令されたデータ・ブロックは、DASD 16の1  
15 回の選択の間に、記録のためにグループ化され  
る。ある状況の下では、データ・ブロックの数は、  
所与のシリンドにおけるトラックの全数より  
少ない数に制限されてよい。例えば、システム・  
ストレージ・アドレス・レジスタ58が8個しか  
20 ない時、最大8個のデータ・ブロックを転送する  
ことができる。キャッシュ・アドレスを即時に利  
用可能とすることから独立して、全ての転送を十  
分に早く処理するように動作するキャッシュ40  
の場合、転送されるデータ・ブロックのグループ  
25 におけるデータ・ブロックの最大数は、シリンド  
におけるトラック数に等しい。データ・ブロック  
がトラックのサブセットである場合(例えば、各  
データ・ブロックがトラックの4分の1を占める  
場合)、SSAR 58が8つであれば、8つのプロ  
30 ックが転送される。同じように、所与のシリンド  
において、2つのトラックでデータの記憶が起  
る。グループ化の要件については、他の順序置換  
方法を採用してよい。そのようなグループ化及び  
書込みは、第10図を参照して説明するように、  
35 タスク指令ステップ190の動作によつて生じ  
る。

これまで説明した走査及びグループ化は、マイ  
クロプロセッサ31Pのかかりの動作時間を必要  
とする。従つて、第9図のステップ281で自由  
なブロックがない時、マイクロプロセッサ31P  
は、カースト・アウト動作へ行かないで、LRU  
リスト46を調べて、キャッシュ40で変更され  
ないで記憶されているデータ・ブロックがあるか  
どうかを調べる。LRUリスト46の中で、部分

(21)

特公 昭 61-43742

41

47 (第1図)にある変更(M)ビットが0のエントリイは、割振可能な(即ち、自由な)ブロックのリストを構成する。第3図のADEBレジスタ76にあるMフィールド124は、そのようなエントリイについてゼロである。従つて、部分47 5の変更ビットが検査される。そして、LRUリスト46の中でMビットがゼロである最初のエントリイが、ホスト11又はDASD16からデータを受取るように割振られる。そのような走査は、LRUリスト46のLRU端部(LRU+0)のところで始まる。LRUリスト46の走査は、制御ストア73(第2図)のJカウンタ37によつて制御される。Jカウンタ37は、それをオール・ゼロへセツトすることによつて301で初期化される(第9図)。Jカウンタ37がゼロの時、LRU 15リスト46を走査しているマイクロプロセッサ31Pは、キャッシュ40で使用時点の最も古いデータ・ブロックを指定するエントリイにアクセスする。ステップ301の初期化の後、ループ300は、LRUリスト46の走査を制御する。先 20ず、302で、LRU+Jのエントリイがマイクロプロセッサ31Pへ読出される。これは、LRUリスト46でINDEXフィールド107を通して指定された登録簿43のエントリイが、システム・ストレージ30からADEBレジスタ76へ 25略読出されることを意味する。303で、登録簿43から読出されたばかりのエントリイが調べられ、Mフィールド124がゼロであるかどうか決定される。もしMフィールド124がゼロであれば、それは変更されていないデータであり、キャッシュ40に記憶されたデータは、DASD16中に記憶された対応するデータに等しい。次に304で、決定されたダイレクトリイ43のエントリイ及びADEBレジスタ76の内容を削除することによつて、キャッシュ40の決定されたばかりの 35アドレス可能記憶領域が、割振りのために解放される。次にループ300は終了し、通路282がとられて、割振ステップ284へ進む。もしMビットがゼロなければ、マイクロプロセッサ31Pは、ループ300を通つて、LRUリスト46の 40次の「使用時点が最も古い」エントリイを指定するように進行する。305で、Jカウンタ37へ1が加えられる。306で、マイクロプロセッサ31Pは、新しいJカウンタの値と、(N+1)

42

の値と比較する。N+1は、第1図の可用閾値49のすぐ上にあるLRUリスト46のエントリイに対応する。MRU-MはLRU+(N+1)と同じ値であることに注意されたい。JカウンタがN+1より小さい時、マイクロプロセッサ31Pはステップ302へ戻り、303のブランチによつてデータが変更されていないことが決定されるか、可用閾値49に達するまで、ループ300が反復される。次に、307で、マイクロプロセッサ31Pは、タスク指名ステップ190へ戻る前に、いくつかの無関連の機能を実行する。ステップ306から190へ戻つた時、キャッシュ・ミスの結果(読出しであつても書込みであつても)、キャッシュ40に記憶されるべきデータについて、 15キャッシュ40内にスペースは割振られていない。所望されたデータ転送は、キャッシュ40中のデータがDASD16へ転送されるまで待たなければならない。データのDASD16への転送は、変更されないブロックを284で割振らせること 20になる。

キャッシュ・ミスが生じた後、又はタスク指名ステップ190の定期的走査中に、キャッシュ40の中で変更されないデータ・ブロックが存在しなければ、記憶システム10は第10図の流れ図に従つてLRUリスト46を走査する。書込み又は読出されるべきデータについて、キャッシュ40中に利用可能なスペースを作るため、自由ブロックを設定する上記の走査は、2つの大きな段階で行なわれる。第1の段階は、放棄リストDL38を調べて、データの放棄が可能かどうかを決定することである。第2の段階は、LRUリスト46が最も使用時点の古いエントリイから可用閾値49まで走査される。この走査でヒット(変更されたデータ)が生じると、そこからグループ化閾 35値57まで第2の走査が始まり、所与のDASD16上に記録されるべきデータ・ブロックの群が設定される。DASD16は上記ヒットによつて決定された、キャッシュ40の変更されているデータを記憶する。

自由なリストがLRUリスト46の1部として含まれている時、キャッシュ・ミスの後の検査はステップ281を省略し、ステップ301を介してループ300へ直接に進む。勿論、299の論理入力、種々の理由で生じてよい。

(22)

特公 昭 61-43742

43

ここで第10図を参照すると、マイクロプロセッサ31Pは論理通路310をたどつて、なされるべき作業を探す。311で、第2図の放棄リスト38がエントリイを有するかどうかを検査される。もし放棄リスト38の中にエントリイがあれば、312で、そのエントリイが取出される。それは、登録簿43のどのエントリイが削除されるべきであるか(ステップ313)を決定し、かつ指標値(ADEBレジスタ76のINDEXフィールド107)をFBLレジスタ79へ転送するためである。これらの動作は、キャッシュ40のために自由ブロックを作る。次に、マイクロプロセッサ31Pは、LRUリスト46の走査のためにループ316及び330へ進むことができる。ステップ314では、LDCBレジスタ62のCCRフィールド95及びMISSフィールド96が1へセットされる論理装置であるかどうか検査される。もしそのような論理装置であれば、マイクロプロセッサ31Pは論理コネクタ319を介して第9図のステップ284へ進む。次いでステップ284から294までが、LDCBレジスタに表示されたCCRフィールド95に関して、記憶システム10の動作を完了する。この時点で、マイクロプロセッサ31PはCCRフィールド95をクリアしない。なぜならば、それは次に受取られた指令を検査して、その指令がチャネル指令再試行の結果として送られたかどうかを決定するからである。

ステップ311で、放棄リスト38にエントリイがないか、314で、CCRフィールドをセットされているLDCBレジスタがなければ、マイクロプロセッサ31PはLRUリスト46を走査するステップへ進む。315で、Jカウンタ37をゼロへセットすることによつて、走査が開始される。次にループ316が開始され、LRUリスト46の最も使用時点が古い指標値から可用閾値49まで走査が実行される。ループ316は、LRU+Jに記憶された指標値によつて指定された登録簿43のエントリイを読出すステップを含む。ループ316を最初に通る時のJカウンタはゼロである。従つて、マイクロプロセッサ31PはLRUリストの使用時点が最も古いエントリイを讀出して、318でMフィールド124を調べる。Mフィールドがゼロである時、320でJカウンタは1だけ増加される。次に321で、Jカ

44

ウンタの値がN+1(LRUリスト46で可用閾値49のすぐ上にあるエントリイ)と比較される。JカウンタがまだN+1に等しくなければ、ループ316が反復される。ループ316が完了すると、論理通路327がとられて、マイクロプロセッサ31Pはタスク指名ステップ190へ戻される。この場合、LRU+0と可用閾値49との間で指定された全てのデータ・ブロックは、変更されたデータを有しない。即ち、第10図に示されたマシン動作は十分な回数だけ実行されたか、他のアルゴリズムによつて、可用閾値49の下にあるエントリイによつて表わされるキャッシュ40の全ての変更されたデータが、DASD16へ転送されている。この場合、キャッシュ40からデータを転送する必要はない。第9図のステップ303で、新しい割振りによつて変更されないデータ・ブロックが発見されるからである。

第10図において、検査されている登録簿43のエントリイのMフィールド124、又はLRUリスト46中の部分(Mビット)47が1である時、走査中に書込ヒットが起つている。次にマイクロプロセッサ31Pは論理通路325をたどり、現在検査されているエントリイで指定された装置について、書込待ち行列(WQ)127へLRU+Jの指標値を転送する。上記の装置は、ADEBレジスタ76のフィールド108にあるDフィールドで指定されている。例えば、もし装置D2がDフィールドに指定されていれば、対応する書込待ち行列127は、ADEBレジスタ76のフィールド107から指標値を受取る。この時点で、装置D2に対応するフラグ・カウンタ127Fは、1だけ増加される。もし書込待ち行列が空であつたならば、フラグ・カウンタは1の値を含む。空でなければ、フラグ・カウンタは書込待ち行列におけるエントリイの数を反映している。この指標値は、キャッシュ40からDASD16へ転送されるべき、最初の指定されたデータ・ブロックを表わす。ヒットの発生はループ316を終了させ、マイクロプロセッサ31Pにループ330を実行させる。

ループ330はLRUリスト46を走査するが、その走査はヒット位置(例えばLRU+1)からグループ化閾値57まで、又は最初に決定されたデータ・ブロックとグループにすることがで

(23)

特公 昭 61-43742

45

きる、適当な数のデータ・ブロックが決定されるまでである。キヤツシユ 40 中でグループにすることが出来るデータ・ブロックは、最初に決定されたデータ・ブロックの CCP 値 (ADEB レジスタ 76 のフィールド 108 参照) に等しい CCP 値を有するデータ・ブロックである。勿論、これらのデータ・ブロックは同一のアドレスされた装置の中にある。

グループ化閾値 57 までの第 2 の走査は、ヒットがループ 316 で生じた時にのみ起る。それは、マイクロプロセッサ 31P が J カウンタを 1 だけ増加させるステップ 331 で始まる。第 1 図の例では、走査されるべき次の LRU エントリイは、LRU+2 に対応する。従つて、登録簿 43 のそのエントリイは ADEB レジスタ 76 へ転送される。次にマイクロプロセッサ 31P は、332 で M フィールド 124 を検査する。その値はゼロに等しいから、マイクロプロセッサ 31P は論理通路 333 をたどつて、ステップ 331 へ戻り、ループ 330 を繰返す。もし M フィールド 124 がゼロでなければ、334 でマイクロプロセッサ 31P は ADEB レジスタ 76 の CCP 値、及び最初に指定されたデータ・ブロック (これは番待ち行列 127 にある) の CCP 値とを検査する。他の実施例としてマイクロプロセッサ 31P の中にある作業レジスタへ、ADEB レジスタ 76 のフィールド 108 で表わされた装置アドレスを含ませ、現在の ADEB レジスタ 76 のエントリイにおける CCP 値と、最初に指定されたデータ・ブロックの CCP 値とを比較するようにしてもよい。注意すべきは、CCP 値が装置上の物理シリンダ・アドレスを表わすことである。こうして、DASD 16 内部の実際のディスク構成は、LRU リスト 46 の走査を変更せず、またそれに影響を及ぼさない。

もしシリンダ・アドレス (CCP 値) が異つていれば、論理通路 333 がとられ、ループ 330 の次の走査が開始される。双方のデータ・ブロックが同一のシリンダにあれば、335 で、マイクロプロセッサ 31P は  $LRU + (J + 1)$  の指標値を番待ち行列 (WQ) 127 へ転送し、フラグ・カウンタ 127F を増加させる。次に 336 で、マイクロプロセッサ 31P は WQ 127 を検査し、CCP 値を有するシリンダについて、8 個

46

のエントリイがあるかどうか決定する。もし 8 個のエントリイがあれば、キヤツシユ 40 から DASD 16 へ転送するためのデータ・ブロックの完全なグループが決定されたことになる。ループ 330 の走査は不完全のままである。従つて、タスク指名ステップ 190 への戻りが生じ、第 11 図を参照して説明する番待ち行列の走査が、グループ化されたデータの転送をスケジュールする。もしグループ化が上限に達していなければ、338 で、マイクロプロセッサ 31P は J カウンタ 37 の内容を検査し、その値を MRU-K (グループ化閾値 57 のすぐ上にある、LRU リスト 46 のエントリイ) と比較する。もし比較が一致すれば、それは第 2 走査の完了を意味し、タスク指名ステップ 190 への通路 337 がとられる。その時点で、転送されるべきデータ・ブロックが 1 個だけであつてもよく、7 個までのブロックがあつてよい。J カウンタの値がまだ MRU-K に達していなければ、通路 333 がとられて、もう 1 回のループ 330 が実行される。

第 11 図は、キヤツシユ 40 に記憶されたデータであつて、DASD 16 へ転送されるべきものを探するため、番待ち行列 127 の全てを走査するマイクロプロセッサ 31P のマシン動作を示す。マシン動作は、タスク指名ステップ 190 から論理通路 340 を介して起動される。341 で、検査されるべき番待ち行列 127 を決定するため、番待ち行列カウンタ (WQK) 127K の内容が検査される。WQK 127K は、DASD 16 の数に等しいディジット位置を有するシフト・レジスタとして処理される。1 つのビットが 1 へセットされており、そのビット位置はどの番待ち行列が検査されるべきかを示す。1 つの番待ち行列の検査が完了すると、1 のビットが次のディジット位置へシフトされ、次の番待ち行列の検査が可能となる。このようにして、ラウンド・ロビン式の待ち行列検査が実行される。

ループ 342 はフラグ・カウンタ 127F を検査して、関連した番待ち行列が空であるか (フラグ・カウンタ 127F がゼロに等しい)、又は非ゼロであつて 1 つ又はそれ以上のエントリイを指示するかどうかを決定する。343 で、番待ち行列カウンタ (WQK) 127K に対応するフラグ・カウンタ 127F が、制御ストア 73 から

(24)

特公 昭 61-43742

47

取出される。フラグ・カウンタ127Fは344で検査される。もしフラグ・カウンタ127Fがゼロであれば、345でWQK127Kが1だけ増進される。346で、シフトされた値が、341で取出された初期値と比較される。もしそれらの値が等しければ、走査は完了している。次にマイクロプロセッサ31Pは、通路347をたどってタスク指名ステップ190へ戻る。比較が一致しなければ、ステップ343-345が反復される。フラグ・カウンタ127Fがゼロでなければ、マイクロプロセッサ31Pは通路348をたどり、350で指定された装置が使用中であるかどうかを検査する。もし装置が使用中であれば、その装置について作業はスケジュールされず、マイクロプロセッサ31Pは論理通路351をたどって、ステップ343へ戻り、ループ342を実行して走査を継続する。350で、もし装置が使用中でなければ、352で、前述した手順を用いてICW連鎖が形成される。353で、形成されたICW連鎖がLKPレジスタ25を介してDACプ  
 20 ログラム56Pへ転送され、そこで実行される。次に354で、キャッシュ40からアドレスされたDASD16へ、実際のデータ転送が起る。355で、本発明と無関連の事後転送動作が起る。356で、第1図のLRUリスト46にあるMビット(部分47)がリセットされ、FBLレジスタ79にあるエントリは、DASD16へ転送されたばかりのデータ・ブロックを含む、キャッシュ40のアドレス可能データ記憶領域の指標値を受取る。これは記憶システム10を準備して、30 FBLレジスタ79に従ってキャッシュ40の中でスペースを割振らせる。357で、LDCBレジスタ62が検査され、CCRフィールド95及びMISSフィールド98が1へセットされているかどうかが決まる。もしキャッシュ・ミスがあれば、358でFBLレジスタ79が読出され、キャッシュ40で解放されたばかりのアドレス可能記憶領域の1つの指標値が得られる。359で、FBLレジスタ79が検査され、それが空であるかどうか決定される。もしそれが空であれば、エラーが生じている。即ち、ステップ358は成功裡に実行されない。次にマイクロプロセッサ31Pは、通路360をたどって、エラー回復手順へ進む。空でなければ、マイクロプロセッサ

48

31Pは、第11図の論理コネクタ361及び第9図の論理コネクタ283を介して第9図の割振ステップ284へ進む。もし357で示されるように、LDCBレジスタのCCRフィールド及びMISSフィールドが1でなかったために、キャッシュ・ミスが生じて、書込待ち行列の走査が起らなかったのであれば、マイクロプロセッサ31Pは、362で本発明と無関係の機能を実行し、タスク指名ステップ190へ戻る。

10 代替方法として、マイクロプロセッサ31Pは、書込待ち行列127の走査を完了するため、ステップ362からループ342へ戻ってよい。しかし、或る設計上の制限が存在する場合、マイクロプロセッサ31Pにとつて、書込待ち行列の中に現われるDASD16への全ての書込みを完了するのではなく、実行されるべき各種の作業を走査するため、タスク指名ステップ190を使用するのが適切であろう。他の設計の制限がある場合には、ループ342を反復するため、直ちに通路  
 15 351へ戻るのがよいかも知れない。

第12図は、例えば第1図のセット・ページ・パラメータ(SPP)指令のようなモード設定型の指令に対する、マイクロプロセッサ31Pのマシン動作を示す。このような指令は、チャネル・アダプタ32を通して受取られる。指令は365で第1図の「アドレス及び指令」評価器50によつてデコードされ、366で、マイクロプロセッサは、その指令がモード設定型の指令であるかどうかを決定する。もしモード設定型の指令でなければ、マイクロプロセッサ31Pは論理通路367をたどって、その指令を実行する。もしモード設定型の指令であれば、368で、マイクロプロセッサ31PはCCW19(第1図)のアドレス・バイト20に関連した適当なLDCBレジスタ62  
 35 にアクセスし、RD(読出し及び放棄)フィールド101をバイト22の値へセットするとともに、SEQフィールド94及び100をセットする。本発明の実施に関する限り、アドレスされた論理装置に対するLDCBレジスタ62の更新は、モード設定指令の実行を完了する。セット・ページ・パラメータ指令は、他の制御パラメータを含んでよく、このような制御パラメータは、記憶システム10の制御データ構造(制御ストア73に置かれるのが望ましい)の中で更新されねばなら



(25)

特公 昭 61-43742

49

ない。369で、マイクロプロセッサ31Pは、通常の如くホスト11へ終了状況信号を送る。マイクロプロセッサ31Pが追加の仕事を探してタスク指名スタック190に戻る前に、370で他の機能を実行してよい。

#### 図面の簡単な説明

第1図はホストへ接続された周辺データ記憶システムを示す、本発明を含む論理図である。更に、本発明を実施するに当つて使用されるチャネル指令、内部指令ワード、及びキャッシュ置換制御リストが示される。第2図は周辺データ記憶システムを制御するためプログラム化されたプロセッサを使用する、第1図のシステムの変更例を示す論理ブロック図である。第3図は第1図及び第2図に示された周辺データ記憶システムで使用される各種のデータ構造を示す図である。第4図はLRUキャッシュ置換(又はオーバーレイ)基準を使用するキャッシュ置換制御リストを示す図である。第5図はチャネル指令ワード及び内部指令ワードを使用する、第1図及び第2図の記憶システムの動作を示す図である。第6図は読出指令を実行するときのマシン動作流れ図、及び第1図及び第2図に示された記憶システムのキャッシュ・ヒット論理フローを示す図である。第7図及び第8図は第1図及び第2図に示される記憶システムにおいてバツキング・ストアから前面ストア(キャッシュ)へシーケンス・データを転送するときのマシン動作流れ図(準備部分及び内部指令ワード部分を含む)である。第9図はキャッシュ・ミスに続く制御処理を示すマシン動作流れ図である。第10図は非同期的書込動作の書込グループ化を示すマシン動作流れ図である。第11図はキャッシュからDASDへ書かれる、グループ化されたデータの待ち行列走査を示すマシン動作流れ図である。第12図はモード設定指令の実行を示す

50

マシン動作流れ図である。

10 ……階層周辺データ記憶システム、11 ……ホスト、12~15 ……I/O接続(チャネル、サブチャネル)、16 ……DASD、19 ……チャネル指令ワード(CCW)、20 ……アドレス・バイト、21 ……指令バイト、22 ……指令修飾バイト、23 ……アドレス・バイト、24 ……内部指令ワード(ICW)、25 ……リンケージ・ポート(LKP)レジスタ、26 ……指令バイト、27 ……指令修飾バイト、28 ……アドレス・バイト、30 ……システム・ストレージ、31 ……制御機構、32 ……チャネル・アダプタ、33 ……データ回路、34 ……装置アダプタ、35 ……装置制御アタッチメント、37 ……Jカウンタ、38 ……放棄リスト(DL)、40 ……キャッシュ、43 ……登録簿、44 ……ハッシュ回路、45 ……分散指標テーブル、46 ……キャッシュ置換LRUリスト、47 ……M(変更)ビット部分、49 ……可用閾値、50 ……「アドレス及び指令」評価器(ACE)、56 ……直接アクセス・コントロール(DAC)、57 ……グループ化閾値、58 ……システム・ストレージ・アドレス・レジスタ(SSAR)、61 ……キャッシュ・アクセス・コントロール(CAC)、62 ……論理装置制御ブロック(LDCB)レジスタ、63 ……チャネル制御ブロック(CCB)レジスタ、69 ……待ち行列(Q)レジスタ、73 ……制御ストア、75 ……他のプログラム、76 ……ADEBレジスタ、77 ……装置バッファ(DEVBUFF)レジスタ、78 ……バツファ・シーケンス・テーブル(BST)、79 ……自由ブロック・リスト(FBL)、31P ……マイクロプロセッサ、50P ……ACEプログラム、56P ……DACプログラム、61P ……CACプログラム。

(26)

特公 昭 61-43742

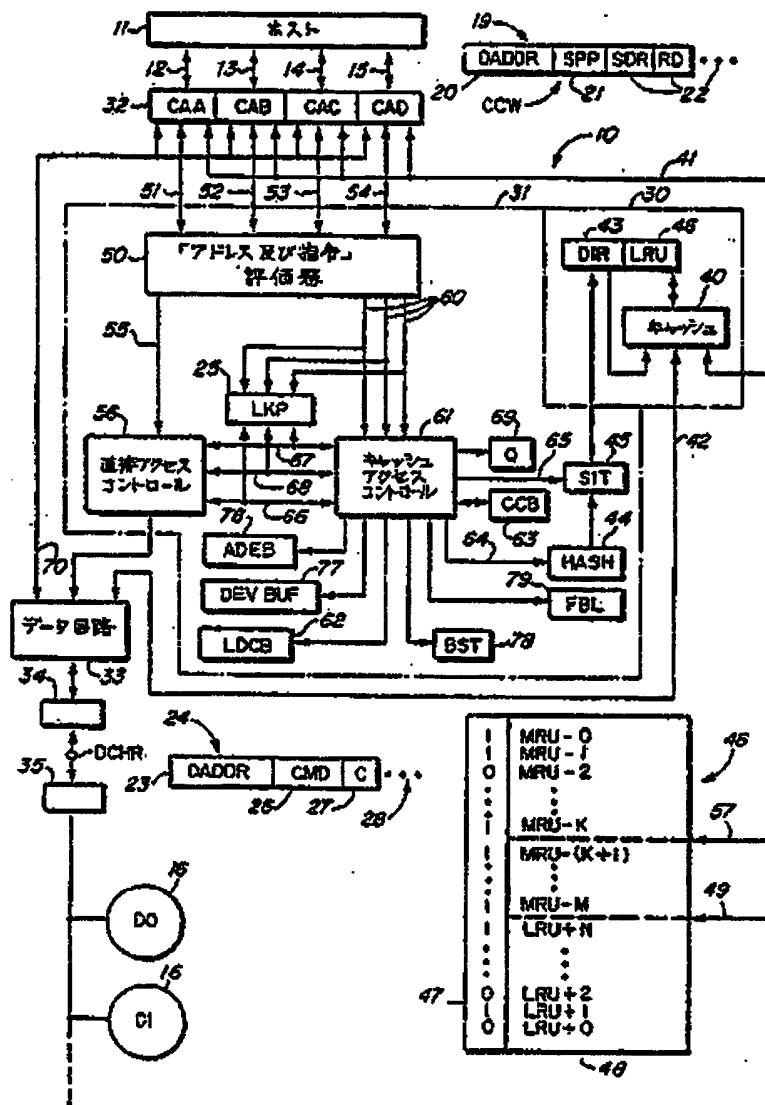
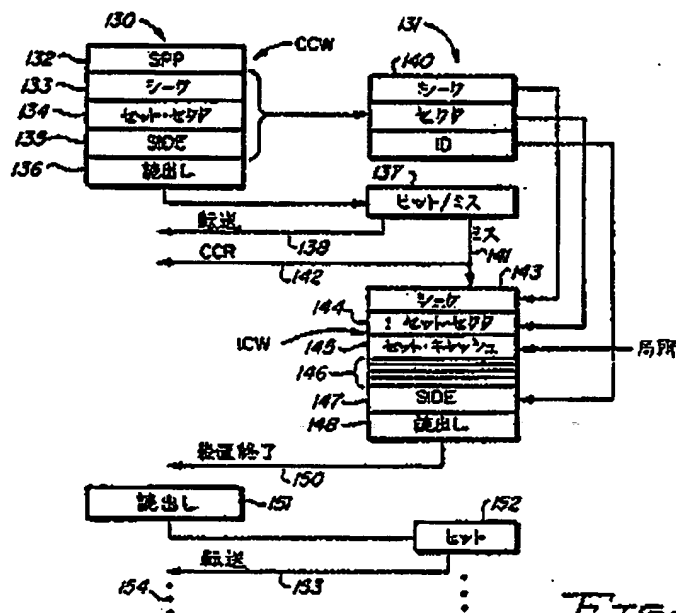
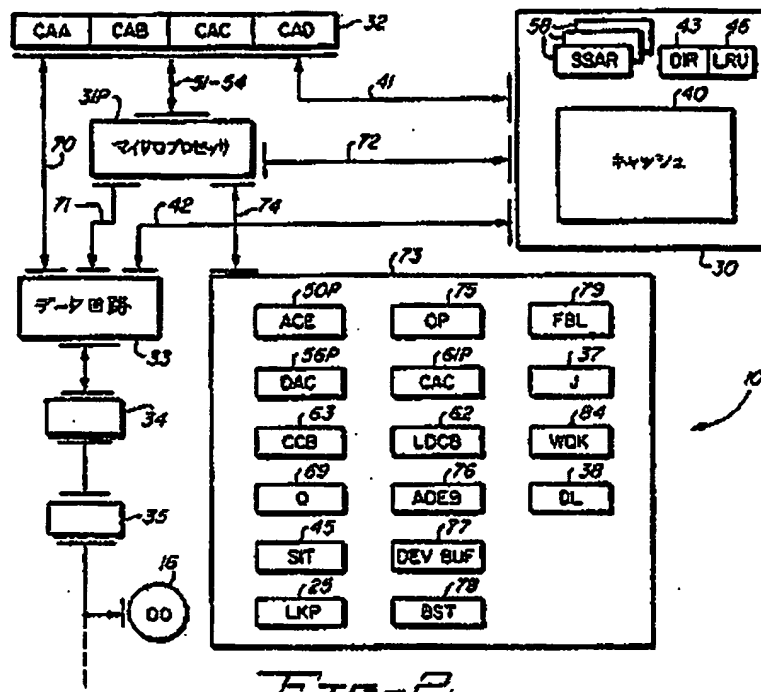


FIG-1

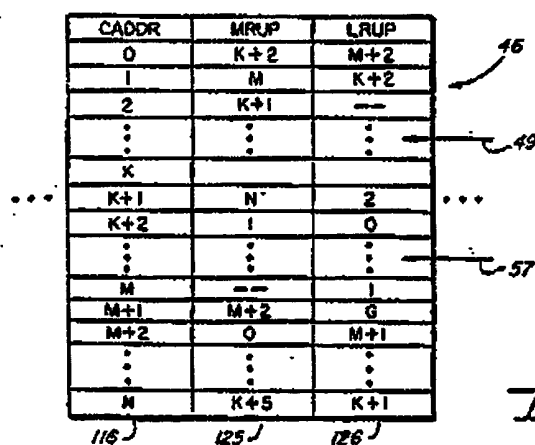
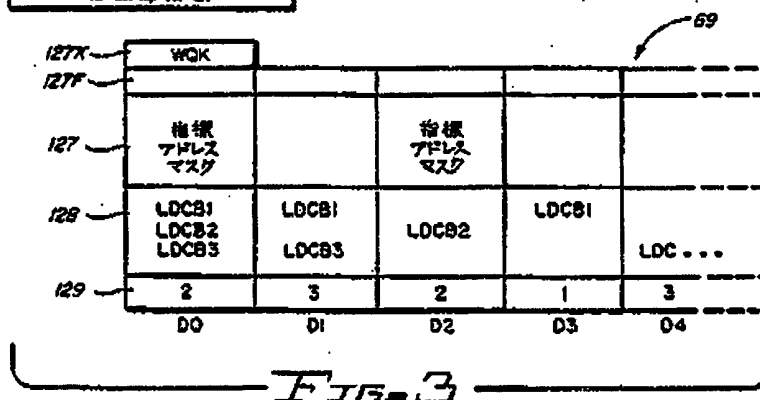
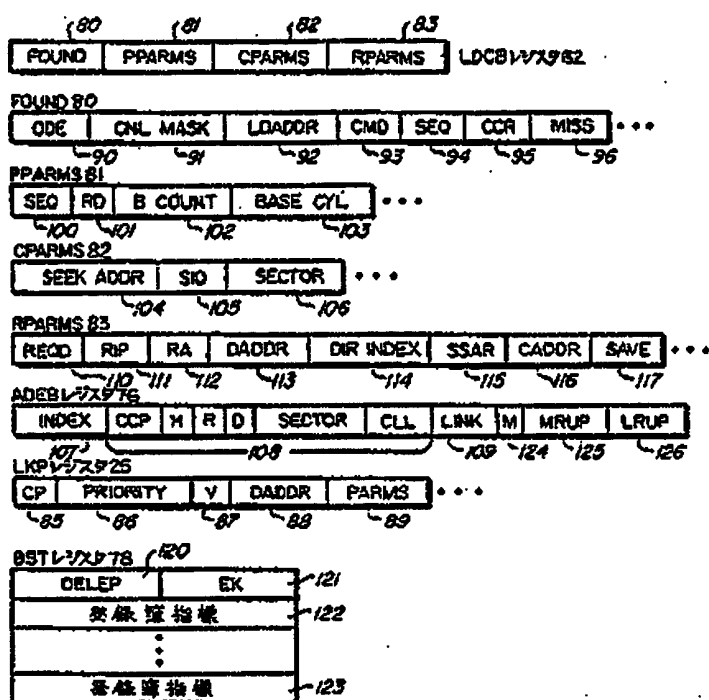
(27)

特公 昭 61-43742



(28)

特公 昭 61-43742





(30)

特公 昭 61-43742

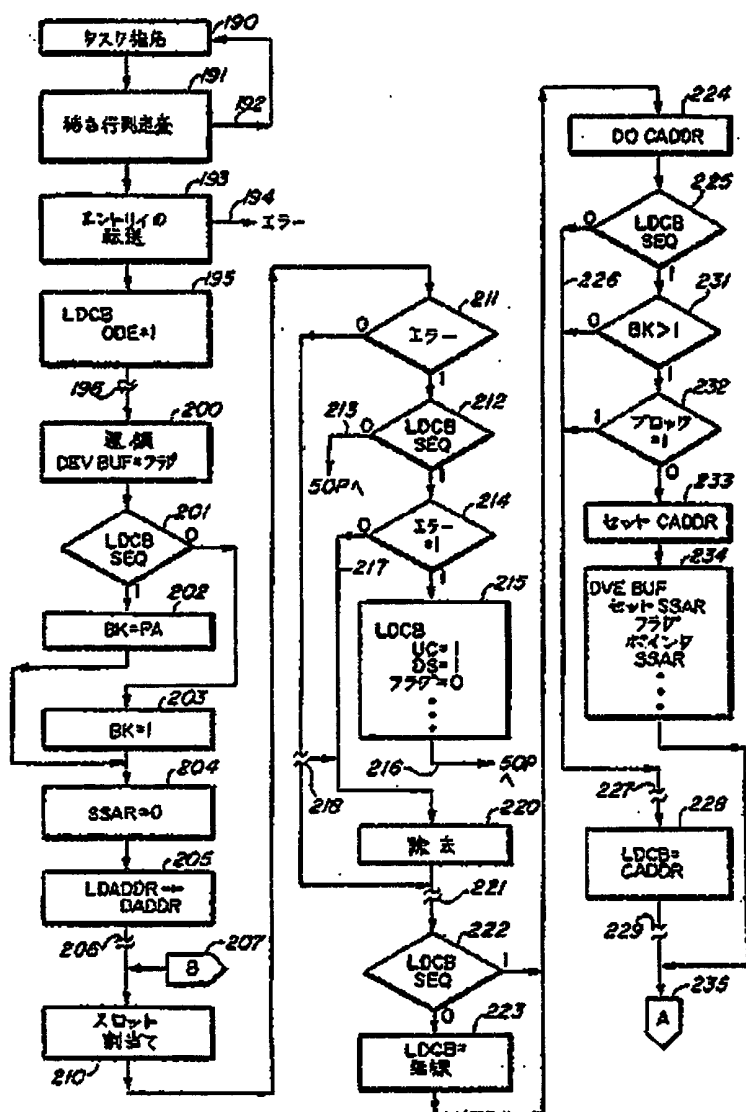


FIG. 7

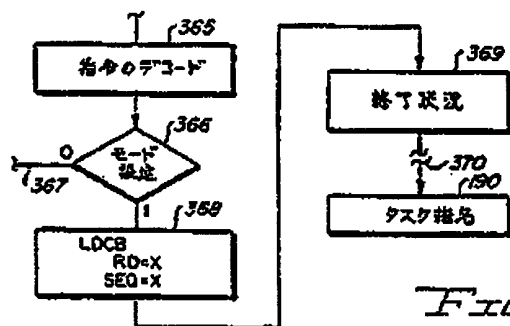


FIG. 12



FIG. 10



(33)

特公 昭 61-43742

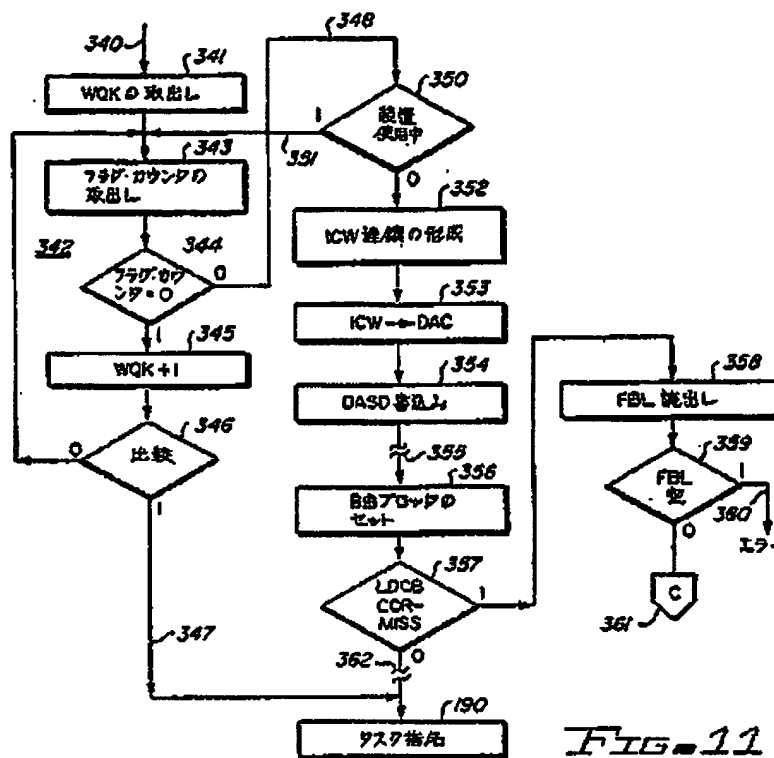


FIG. 11

**THIS PAGE BLANK (USPTO)**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)